1-1-2015

Therapeutic Applications and Specificity of Action of Designer Nucleases for Precision Genome Engineering

Chukwuka Anthony Didigu
University of Pennsylvania, chuka.didigu@gmail.com

Therapeutic Applications and Specificity of Action of Designer Nucleases for Precision Genome Engineering

Abstract
Designer nucleases allow for the precise modification of a given DNA sequence by the introduction of a sequence-specific double strand break. This targeted genetic engineering confers the ability to modify genomes of complex organisms, and has far-reaching applications in human medicine, agriculture, and biotechnology. As these nucleases act in a sequence specific manner, understanding their specificity is of paramount importance to prevent potentially genotoxic side effects. In this thesis, I assessed the ability of a class of designer nucleases (ZFNs)--zinc finger nucleases--to simultaneously inactivate two genes encoding entry factors required for HIV infection in human CD4 T cells. Additionally, I sought to develop a high-throughput means of identifying sites of designer nuclease off-target activity across the genome, in an effort to better understand the factors governing designer nuclease specificity. This work demonstrates the ability of ZFNs to simultaneously modify two distinct genetic loci in primary human CD4 T cells--the main target of HIV infection. These gene-modified cells are protected from HIV infection and represent a novel means of treating--and potentially curing HIV infection. This work also demonstrates that DNA double-strand breaks introduced by a single designer nuclease at on- and off-target loci can result in the formation of genomic rearrangements. Taken together, this work advances in the field of genome engineering on two fronts--a novel therapeutic application of designer nucleases and a novel means of detecting off-target genomic modification.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Cell & Molecular Biology

First Advisor
Robert W. Doms

Second Advisor
John E. Wherry

Keywords
CRISPR, Designer nuclease, Genome editing, HIV, TALEN, Zinc Finger nuclease

Subject Categories
Biology | Biomedical | Microbiology

THERAPEUTIC APPLICATIONS AND SPECIFICITY OF ACTION OF DESIGNER NUCLEASES FOR PRECISION GENOME ENGINEERING

Chukwuka Anthony Didigu

A DISSERTATION

In Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

In Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

_____

Robert W. Doms, Professor of Pathology and Laboratory Medicine

Graduate Group Chairperson

_____

Dan Kessler, Associate Professor of Cell and Developmental Biology

Dissertation Committee

John E. Wherry (Chair), Associate professor of Microbiology

Beatrice Hahn, Professor of Medicine and Microbiology

Carl H. June, Professor of Pathology and Laboratory Medicine

David B. Roth, Professor of Pathology and Laboratory Medicine

THERAPEUTIC APPLICATIONS AND SPECIFICITY OF ACTION OF DESIGNER NUCLEASES FOR PRECISION GENOME ENGINEERING

DEDICATION

There are a number of people I have to thank, for their tireless support over the last few years. I would like to thank my family, for their understanding. I would like to thank the University of Pennsylvania MSTP—particularly Maggie Krall—for always knowing the right thing to do. And I would like to thank my friends—especially Dr Gregory Charak—for never letting me forget what really matters. Finally I would like to thank my mentor, Dr Robert W Doms. His mentoring style, keen intellect, logical thought process, and warmth are traits that drew me to his lab and that made the last three years the single most instructive experience I have ever had. To him, I am truly grateful.

ACKNOWLEDGMENT

This work would not have been possible without the efforts of a number of people both within and outside of the University of Pennsylvania. As such I would like to thank the following people for their immeasurable contributions to this thesis—Dr Craig Wilen, Dr Carl June, Dr Frederick bushman, Dr James Riley, Dr Andrew Stern, Nirav Malani, Mary Jane Drake, Dr Paul Bates, Dr Nicholas Parrish, Dr Beatrice Hahn, Dr Jason Wojcechowskyj, Zahra Parker, and Val hardy.

ABSTRACT

THERAPEUTIC APPLICATIONS AND SPECIFICITY OF ACTION OF DESIGNER NUCLEASES FOR PRECISION GENOME ENGINEERING

Chukwuka A. Didigu

Robert W. Doms

Designer nucleases allow for the precise modification of a given DNA sequence by the introduction of a sequence-specific double strand break. This targeted genetic engineering confers the ability to modify genomes of complex organisms, and has far-reaching applications in human medicine, agriculture, and biotechnology. As these nucleases act in a sequence specific manner, understanding their specificity is of paramount importance to prevent potentially genotoxic side effects. In this thesis, I assessed the ability of a class of designer nucleases (ZFNs)—zinc finger nucleases—to simultaneously inactivate two genes encoding entry factors required for HIV infection in human CD4 T cells. Additionally, I sought to develop a high-throughput means of identifying sites of designer nuclease off-target activity across the genome, in an effort to better understand the factors governing designer nuclease specificity. This work demonstrates the ability of ZFNs to simultaneously modify two distinct genetic loci in primary human CD4 T cells—the main target of HIV infection. These gene-modified cells are protected from HIV infection and represent a novel means of treating—and potentially curing HIV infection. This work also demonstrates that DNA double-strand breaks introduced by a single designer nuclease at on- and off-target loci can result in the formation of genomic rearrangements. Taken

together, this work advances in the field of genome engineering on two fronts—a novel therapeutic application of designer nucleases and a novel means of detecting off-target genomic modification.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Chapter 1. Designer nucleases are a flexible tool for precision genome engineering

I. Genome engineering using designer nucleases

The ability to introduce stable, heritable, genetic modifications in complex organisms provides a unique opportunity to alter such organisms for therapeutic, biotechnology and research purposes. This process of genome engineering can be achieved using a diverse family of molecules called designer nucleases. Designer nucleases are chimeric molecules that can be engineered to bind a given DNA sequence and once bound to their target, introduce a double-strand break (DSB). This DSB serves as the nexus for a number of genomic alterations including inactivation of target genes, gene correction, and the insertion of new coding sequences (reviewed in (1)). Inactivation of a target gene requires DSB repair to occur via the error-prone cellular non-homologous end-joining (NHEJ) pathway, which most often generates random insertions and deletions that result in a non-functional gene product (2). Conversely, the presence of a DNA template with homology to the region surrounding the DSB—such as a sister chromatid—allows for DSB repair to occur via the higher fidelity homologous repair pathway. As such, the concomitant introduction of a DSB by a designer nuclease and delivery of a custom DNA template with homology to the site of the DSB allows for the introduction of an investigator-designed DNA sequence into a given locus. These sequences can range from single base pair substitutions to the insertion of large segments of coding DNA. To date, the three most widely

used designer nucleases include **Z**inc **F**inger **N**ucleases (ZFNs), **T**ranscription **A**ctivator-**L**ike **E**ffector **N**ucleases (TALENs) and **C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeat-associated nucleases (the CRISPR/Cas system).

II. Zinc Finger Nucleases

ZFNs consist of a DNA-binding Cys_2-His_2 zinc finger protein (ZFP) fused to the nuclease domain of the dimeric FokI restriction endonuclease (3-5). The Cys_2-His_2 zinc finger domain is one of the most common DNA binding motifs found in eukaryotic transcription factors. These domains have a ββα-fold conformation, and amino acids in the α-helix of the zinc finger contact 3-4 base pairs in the major groove of DNA. The ability to target longer sequences was made possible by the development a polypeptide linker that allows for the generation of tandem arrays of zinc fingers, and has resulted in enhanced ZFN specificity (6). To date, ZFPs that bind most of the 64 possible DNA triplets have been designed, greatly aiding the synthesis of ZFNs that target sequences long enough to ensure genome-wide specificity.

ZFNs function as a pair, and to introduce a double stranded break at a given locus, members of a ZFN pair must first bind their target DNA sequence, with each half of the pair binding to an opposite strand of DNA. This binding event brings the members of a ZFN pair into close enough proximity to allow for the formation of a catalytically active nuclease complex following dimerization of the nuclease domains, which then cleaves the underlying DNA sequence. ZFN

binding is highly specific, as cleavage will only occur if the ZFNs recognize and bind to an appropriate sequence, and the members of the ZFN pair bind in the correct orientation with a spacer requirement of 5-6 base pairs of DNA between each ZFP target sequence. Additionally, the FokI nuclease domain which normally functions as a homodimer, has recently been modified to function as a heterodimer by introducing charged residues at the nuclease dimer interface and placing these positively and negatively charged residues onto opposite sides of the monomers that comprise the FokI dimer(7). As two positively or negatively charged FokI monomers will not interact, these alterations allow the FokI nuclease to function as an obligate heterodimer, eliminating the possibility of one-half of a ZFN pair binding on opposite DNA strands at an off-target site and introducing an unwanted DSB(7).

Despite their ability to target a wide array of sequences, ZFNs can be quite challenging to design owing in large part to the context-dependent nature of their DNA binding domains i.e. DNA triplet binding preference of each zinc finger domain in a ZFP array is influenced by any ZFPs to which it is directly linked. As such, early ZFP arrays used to design ZFNs displayed a high failure rate with many ZFNs failing to bind and cleave their target sequences. Recent improvements in the ZFN assembly platforms have made ZFN design somewhat easier, but these issues of context dependence and the relatively high cost of making ZFNs has led many groups to explore alternative platforms to accomplish precision genome editing.

III. Transcription Activator-Like Effector Nucleases

TALENs bear a similar overall architecture to ZFNs in that they also function as a pair, with a DNA-binding protein domain fused to the nuclease domain of the FokI endonuclease. The TALEN DNA binding domain is based on the transcription activator-like (TAL) DNA-binding proteins found in *Xanthomonas* bacteria. Unlike ZFPs, the DNA binding of TAL effectors is almost entirely free of context dependent effects. Each TAL repeat usually consists of 33-35 amino acids, where all the residues are highly conserved except for residues in the 12 and 13 position. Together these two amino acids are termed the repeat-variable diresidue (RVD) and they confer the nucleotide specificity of TAL effectors. Specifically, the 13th residue contacts a single nucleotide in the major groove of DNA while the 12th residue stabilizes the short loop structure formed by the RVD (8,9). As each RVD usually binds a single nucleotide, modular assembly of TAL repeats can be used to generate TAL protein arrays capable of binding any DNA sequence as long as that sequence begins with a Thymine base. While TALENs can also be used with obligate heterodimeric FokI nuclease domains, and can target long sequences to achieve genome-wide specificity as is done with ZFNs, they have a more variable spacer requirement than ZFNs, usually in the range of 10-20bp and this could, in theory, contribute to an increased rate of off-target binding activity. Additionally, the repetitive nature of the TAL arrays makes them difficult to synthesize, and so while TALENs lack the context-dependent effects seen with ZFNs, their variable spacer lengths and challenging synthesis have hindered their widespread adoption.

IV. The CRISPR/Cas system

The CRISPR/Cas system is the only known form of adaptive immunity against phage infection in prokaryotes, and is made up of two main components—arrays of **c**lustered **r**egularly **i**nterspaced **s**hort **p**alindromic **r**epeats (CRISPRs), and **C**RISPR **as**sociated (Cas) genes that are usually located adjacent to the CRISPR arrays. Following phage infection, the CRISPR/Cas system functions by incorporating phage DNA into the host genome, synthesizing CRISPR RNAs (crRNA) with complementarity to the incorporated/invading phage DNA, forming crRNA-phage DNA hybrids, and degrading these hybrids through the nuclease activity of the Cas proteins (reviewed in (10)). Despite the diversity of bacteria and archaea with naturally occurring CRISPR/Cas systems, these systems all have the same overall architecture. CRISPR loci contain arrays of short nucleotide repeats separated by non-repeating spacer sequences. These spacer sequences, which are derived from invading phage, are called protospacers, and are flanked by short DNA sequences called protospacer adjacent motifs (PAMs). There exist three main CRISPR/Cas systems—Type I, II and III, and of these the type II CRISPR/Cas system is arguably the most well studied. In this system, transcription at a CRISPR locus generates a pre-crRNA, which is then processed by Cas proteins into a crRNA containing, among other things, the phage derived protospacer sequence. This crRNA then binds a trans-activating crRNA (tracrRNA) and this complex guides sequence specific binding of the crRNA to the invading phage DNA and subsequent cleavage of the resulting RNA-DNA

hybrid by the Cas9 nuclease. Several recent studies have recreated the CRISPR/Cas system by generating a hybrid RNA—generally termed a guide RNA or gRNA(11-13). This gRNA has both complementarity to a target sequence located adjacent to a PAM sequence (similar to the crRNA), and a stem loop structure (similar to that present in the tracrRNA) that recruits the Cas nuclease (Cas9) to the site of the RNA-DNA complex. As such, co-delivery of a gRNA and a Cas9 encoding plasmid into a eukaryotic cell will result in the introduction of a DSB in a sequence-specific manner at any desired locus adjacent to a PAM sequence. A huge advantage of the CRISPR/Cas system is the ease with which gRNAs can be designed and used to effectively target a diverse array of sequences. Additionally, the gRNA is a relatively short oligonucleotide (usually ~ 100bp), thus CRISPR/Cas gene targeting is generally more affordable than ZFN or TALEN mediated gene targeting. However, as CRISPRs do not have any of the specificity requirements built into the ZFN and TALEN architecture such as the binding requirement on opposite strands or the use of obligate heterodimeric nuclease domains, there remain concerns regarding their genome-wide specificity.

V. Project goal I - Therapeutic applications of ZFNs to treat HIV

To date, ZFNs, TALENs and CRISPR/Cas nucleases have been used to effectively target diverse genetic loci in cell lines and primary cells (including stem cells) (5,14-23). One of the earliest therapeutic applications of designer nucleases was the inactivation of a gene required for HIV infection in HIV-

susceptible CD4 T cells (14). To infect a cell, HIV must first bind to its primary receptor CD4, and then to one of two coreceptors—CCR5 or CXCR4. As individuals with a naturally occurring inactivating mutation in *ccr5* are highly resistant to HIV infection, this study attempted to replicate this phenotype by inactivating the *ccr5* gene using ZFNs and showed that these gene-modified cells were both resistant to HIV infection and were selected for in the presence of viruses that use CCR5 to enter cells (14). This promising preclinical data led to one of the first human clinical trials involving a designer nuclease—the study of the autologous transfer of CCR5-ZFN modified CD4 T cells in HIV infected individuals (www.clinicaltrials.gov; Trial identifiers NCT00842634, NCT01044654, NCT01543152). While this trial holds great promise for a novel HIV treatment and a potential cure, if successful it would likely be most effective in patients that harbor viruses that exclusively use CCR5 to infect cells. As up to half of HIV-infected individuals harbor viruses capable of using CXCR4 to infect cells (24), I sought to make this therapeutic application of ZFNs more broadly applicable in the first part of my thesis by using ZFNs to simultaneously inactivate both *ccr5* and *cxcr4* in primary human CD4 T cells (Chapters 2 and 3).

VI. Project goal II – Identification and characterization of sites of genome-wide designer nuclease off-target activity

Although designer nucleases have the potential to change the way we treat diseases and perform research, we do not yet have a complete understanding of the genome-wide specificity of these nucleases or all the

factors that influence cleavage activity at unwanted sites. A number of studies have shown that while each designer nuclease exhibits a high level of cleavage activity at their target site, there is some level of unwanted cleavage activity at off-target sites. By searching the genome for sites with sequence similarity to the nuclease target site and analyzing these potential off-target sites for evidence of nuclease activity, several studies have identified some sites of off-target activity(14,15,25). However, while sequence similarity is likely a predictor of off-target activity, there is mounting evidence suggesting that it is not the sole predictor of off-target activity (26). Additionally, there are currently no studies comparing the specificity of the three main classes of designer nucleases. Hence as part of my thesis, I attempted to identify sites of off-target activity for the three main classes of designer nuclease using a novel high-throughput assay to detect genome-wide translocations involving nuclease-induced double strand breaks (Chapter 4).

Chapter 2. Searching for a cure: HIV infection and novel treatment strategies

I. The HIV epidemic: A persistent problem

Since its identification as the etiologic agent of acquired immune deficiency syndrome (AIDS), human immunodeficiency virus (HIV) has claimed the lives of millions around the world. The introduction of antiretroviral therapy (ART) for HIV infection has dramatically altered the course of the HIV pandemic by delaying the progression to AIDS and increasing the lifespan of infected individuals (27-30), while decreasing HIV transmission rates (31-34). However, the prospect of eradication remains a distant one, as there are still 34 million people living with HIV worldwide, and more than 2 million new infections every year. Additionally, as the HIV infected population ages, we are learning that control of HIV infection by ART does not completely restore health (reviewed in (35)). Patients on ART remain at increased risk of developing cardiovascular disease compared to age-matched uninfected controls (36), and HIV infection also significantly increases the risk of developing kidney disease, osteoporosis, and a number of non-AIDS defining malignancies, even when patients are well controlled on ART (37-40). Moreover, ART itself is not without side effects, as certain classes of antiretroviral drugs directly contribute to some of these comorbidities. These HIV-associated morbidities that persist despite ART, the high cost of ART, and the requirement for daily adherence to ART to control

infection, all highlight the pressing need for strategies to achieve drug-free control and an eventual cure for HIV.

II. Barriers to a cure: The HIV reservoir

The major obstacle to achieving a cure stems from the fact that HIV not only infects and kills cells involved in combating infection (CD4 T cells), but also establishes a stable reservoir in these cells that goes undetected by the immune system (41-44). This latent reservoir is present in resting memory CD4 T cells, is largely unaffected by ART, and is responsible for the rapid rebound in viremia following ART cessation (43-47). As such, elimination of the reservoir is likely required for a cure. Based on our understanding of HIV latency, it has long been thought that reactivation of the latent reservoir would result in the detection of infected cells and their subsequent destruction by the immune system. The pursuit of this "shock-and-kill" approach has led to the identification of pharmacologic compounds capable of reactivating latent HIV *in vitro* (reviewed in (48)), and there is evidence that one such drug, the histone deacetylase inhibitor vorinostat, is capable of disrupting HIV-1 latency to some degree in HIV infected individuals on ART (49). However, a recent study by Shan et al showed that following drug-induced reactivation of latency in CD4 T cells from patients on ART, these cells were not killed by autologous cytotoxic lymphocytes (CTLs), due to defects in the quality of the HIV-specific CTL response in ART-experienced patients (50). These results suggest that pharmacologic reactivation of latency alone may not be enough to eliminate the reservoir. As such, there is a

need for alternative or complementary approaches that will enhance the ability of the immune system of HIV-infected persons to identify and kill cells harboring reactivated viruses if we hope to effectively eliminate the HIV reservoir using this strategy. A recent landmark study by Ho et al has also called into question our previous estimates of the HIV reservoir, and the ease with which the reservoir can be reactivated (51). This work stemmed from the observation that DNA measurements of the HIV reservoir size are up to 2-logs higher than measurements obtained using the standard viral outgrowth assay—an assay that estimates the reservoir size by measuring HIV production following maximum *in vitro* T cell activation of resting CD4 T cells from HIV-infected individuals (52). This difference in reservoir size estimates was initially believed to represent defective proviruses—a belief supported by the error-prone nature of HIV replication (53-58). However, a careful characterization of these uninduced proviruses by Ho and colleagues revealed that up to 12% of them are actually genetically intact, integrate into active sites of transcription, and when synthesized, display replication kinetics comparable to those of latent viruses induced by T cells activation. This study has 2 major implications for the shock-and-kill approach and the HIV cure field in general—the replication-competent HIV reservoir is considerably larger than previously thought, and reactivation of latent HIV is not determined solely by the activation state of a T cell, but may in fact be in part a stochastic process.

III. Gene therapy for a cure

Genetic manipulation of long-lived primary CD4 T cells and hematopoietic stem cells (HSCs) to prevent HIV infection has long been viewed as a viable means of achieving ART-free control of infection, and following the recent report of a cure for HIV (59,60), there has been a surge of interest in exploring gene therapy-based approaches to treat HIV. This cure was achieved following an allogeneic hematopoietic stem cell transplant (HSCT) to treat leukemia in an HIV-infected man, and was performed using cells from a donor bearing an inactivating mutation in both copies of the C-C chemockine receptor type 5 (CCR5) gene—whose protein product is the primary HIV entry coreceptor (59,60). This mutation—known as *ccr5*Δ32—confers resistance to HIV infection in homozygotes, and delays the progression to AIDS in heterozygotes (61-63). Following the transplant, the patient was taken off ART and in the ensuing years has remained free of HIV, with undetectable viral loads and substantial decreases in HIV-specific antibodies, suggesting that a cure has indeed been achieved. However, there remains a great deal of speculation regarding the reason for the cure. One possible explanation is that the complete donor chimerism achieved following allogeneic transplantation of CCR5-negative cells simply created an environment incapable of supporting HIV infection by eradicating all cells susceptible to infection with CCR5-using HIV. However, this may not fully explain the observed cure as the patient in question also had low levels of HIV capable of entering cells in a CCR5-independent manner by using the other HIV coreceptor C-X-C chemokine receptor type 4 (CXCR4), and yet

these viruses failed to expand following the transplant. That such viruses can expand in vivo in the face of selective pressure against CCR5 is evidenced by the fact that the most common cause of virologic failure following treatment with the CCR5 antagonist maraviroc is outgrowth of pre-existing CXCR4-using HIV strains (64,65). As such, alternative explanations for the cure including the role of graft-vs-host-disease (GVHD) in clearing the infection have been considered, with the assumption that the development of GVHD following the transplant led to the detection and donor cell-mediated clearance of all host immune cells including those cells harboring the latent HIV reservoir. Another potential reason for this remarkable cure is the destruction of the HIV reservoir by the conditioning chemotherapy and total body irradiation administered prior to the transplant. While such transplants have been performed in HIV infected patients in the past with no effect on their HIV infection, recent evidence suggests that allogeneic stem cell transplants with CCR5-positive cells may in fact have an effect on the size of the HIV reservoir as measured by the viral outgrowth assay(66).

The striking resistance to HIV infection observed in *ccr5*Δ32 homozygotes and the recent report of a cure following HSCT using *ccr5*Δ32 cells has spurred several gene therapy efforts to block HIV infection at the level of entry in an attempt to reproduce this HIV-resistant phenotype in individuals lacking the delta32 mutation. As a result, several promising preclinical studies have successfully provided some form of protection against HIV by targeting infection at the level of entry and a number of these studies have been advanced to human clinical trials (Table 1).

IV. HIV Entry: The Basics

The HIV life cycle begins with entry of the virus into susceptible cells. To enter a cell, HIV must first bind to its primary receptor CD4, and then to one of two coreceptors—CCR5 or CXCR4 (Figure 1) [5-13]. The choice of coreceptor used by the virus is intimately linked to disease acquisition and pathogenesis as the majority of transmitted viruses use CCR5 to enter cells [14-17], while the appearance of viruses capable of using CXCR4 during infection is associated with a more rapid progression to AIDS [18-21].

Entry of HIV into target cells is mediated by the type I integral membrane viral glycoprotein Env. Env is synthesized as a polypeptide precursor termed gp160 which undergoes several of modifications within the cell as it is transported to the cell surface, including extensive N-linked glycosylation, and cleavage by cellular proteases into the extracellular gp120 and the membrane-spanning gp41 subunits [27-29]. The extensive glycosylation of Env contributes to its ability to evade humoral immune responses, as the carbohydrate moieties covering its surface are poorly immunogenic and may shield potentially antigenic epitopes on the glycoprotein from recognition by the immune system [30-32]. Following gp160 cleavage—an event required for subsequent membrane fusion—gp120 and gp41 maintain their association via non-covalent interactions and are transported to the cell surface where they exist as a trimer of heterodimers that is ultimately incorporated into the viral membrane. Prior to CD4

binding—the first essential step in HIV entry—a number of cell-surface molecules are capable of mediating Env-dependent attachment of the virion to target cells. One such attachment factor is the C-type lectin CD209 or dendritic cell-specific intercellular adhesion molecule (ICAM) grabbing non-integrin (DC-SIGN). Expressed on dendritic cells, DC-SIGN and several other lectins are capable of binding Env and boosting infection *in vitro* by facilitating *trans*-infection of surrounding CD4 T cells by dendritic cell-bound virions (reviewed in [33]). More recently, monomeric gp120 from some HIV strains has been shown to bind the gut homing integrin α4ß7 [34,35], which is expressed on activated CD4 T cells. This finding is of particular interest as the depletion of CD4 T cells in the gut-associated lymphoid tissue (GALT) early in infection is a hallmark of HIV disease [36-38] and ß7 integrins mediate trafficking of lymphocytes to the gut mucosa [39]. However, it is not clear whether α4ß7 supports binding of trimeric Env on the surface of virions, and thus its relevance to the gut pathology associated with HIV infection remains unknown. Furthermore, while interactions with these and other attachment factors influence HIV infection *in vitro,* little is known regarding their significance *in vivo*.

The gp120 subunit of Env is composed of five relatively conserved (C1 – C5) and five more variable (V1 – V5) regions [40]. The conserved regions form the proximal core of gp120 while intrachain disulfide bonds in the variable regions of gp120 result in the formation of five variable ‘loop’ structures that make up the most exterior portion of the gp120 ectodomain [41]. The CD4 binding site on gp120 is a well conserved cavity formed at the interface of the

inner and outer domains of the glycoprotein [42]. Following binding of gp120 to CD4, a series of conformational changes occur including the rearrangement of two pairs of ß-sheets from the gp120 inner and outer domains that come together to form a four-stranded ß-sheet structure termed the bridging sheet. The bridging sheet links the inner and outer domains of gp120 and interacts with the viral coreceptor, be it CCR5 or CXCR4 [42,43]. gp120 binding to CD4 also results in enhanced exposure and reorientation of the V1/V2 and V3 loops of gp120, outward rotation of each gp120 monomer to reveal the gp41 stalk, and hinge-like movements in CD4 that are thought to bring the viral membrane in close proximity to the target cell [42,44,45]. Together, all of these events culminate in the creation and exposure of the coreceptor-binding site.

The HIV coreceptors belong to the family of chemokine receptors—seven-transmembrane G-protein coupled receptors with prominent roles in immune cell trafficking. These receptors have three extracellular and intracellular loops, extracellular N-termini, and intracellular C-termini. While several different chemokine receptors are capable of mediating HIV entry *in vitro,* current evidence suggests that the CCR5 and CXCR4 chemokine receptors are the most frequently utilized *in vivo* [46-48]. Viruses capable of utilizing CCR5 alone, CXCR4 alone, or both coreceptors, are labeled R5-tropic, X4-tropic, and RX54 or dual-tropic viruses respectively. CCR5 is the primary coreceptor for the majority of HIV-1 isolates and is expressed on CD4 T-cell subsets, macrophages and dendritic cells, while CXCR4 is less commonly used, but is expressed on a wide variety of cells both within and outside the immune system [49]. For reasons that

remain unclear, the majority of transmitted viruses utilize CCR5 irrespective of the route of transmission and despite the availability of target cells expressing CXCR4 [17,50,51]. Multiple lines of evidence including mutational analyses, studies of small molecule inhibitors, and inhibition by coreceptor-specific blocking antibodies suggest that both the second extracellular loop (ECL2) and sulfated tyrosines within the N-terminus of the coreceptors interact with the V3 loop of gp120 and mediate coreceptor binding [52-57]. The V3 loop is also known to be a key determinant of coreceptor preference as the presence of positively charged amino acids at positions 11 and or 24/25 of V3 is correlated with CXCR4 usage [58,59].

The HIV fusion machinery is contained within the gp41 subunit of Env, which is comprised of a large cytoplasmic domain, a membrane-spanning segment, and an ectodomain that maintains contact with gp120. The ectodomain contains a typical fusion peptide—a stretch of hydrophobic amino acids at the N-terminus [60,61]—along with two α-helical heptad repeats (HR), the N-terminal HR1 and the C-terminal HR2 repeats [62,63]. The current model of gp41-mediated fusion is based on studies performed using HIV fusion inhibitors, crystal structures, and structural similarities between gp41 and other well-characterized type I membrane fusion proteins including the influenza virus glycoprotein, hemagglutinin (HA) [63-66]. In this model, the sequential interaction of Env with CD4 and a coreceptor results in exposure of the fusion peptide, which then inserts into the plasma membrane of the host cell, causing gp41 to physically link both membranes. Subsequently, the three HR1 domains of the

Env trimer interact with one another to form a coiled coil, and the three HR2 segments fold back on the HR1 trimer creating a six-helix bundle that brings the viral and host cell membranes in close contact with one another, allowing for mixing of the two membranes and formation of the fusion pore. Fusion between Env and the host cell was long thought to occur at the plasma membrane as HIV entry occurs in a pH-independent manner [67] and Env is capable of mediating fusion between neighboring cells, provided that they express CD4 and an appropriate coreceptor [68]. However, recent work using *trans* dominant-negative mutants of proteins involved in clathrin-mediated endocytosis [69] along with elegant studies using single-virion imaging [70] have demonstrated a clear role for components of the endocytic pathway in HIV entry in a number of cell lines. These studies were performed using immortalized cell lines, and as such, the role of the endocytic pathway in HIV entry into relevant cell types *in vivo* is yet to be determined.

The cure achieved following allogeneic HSCT using *ccr5*Δ32 cells highlights how our understanding of a very basic question—how a virus enters its host cell—has led to the development of new antiviral drugs and therapeutic approaches that have brought us a step closer to controlling the global HIV pandemic.

VI. Pharmacologic inhibition of HIV entry

The multi-step process by which HIV enters cells provides a series of unique targets for interventions to prevent viral entry including receptor and coreceptor binding, and membrane fusion. Efforts to inhibit these steps have led to the discovery of a new class of anti-HIV drugs—the HIV entry inhibitors (reviewed in [71]). A number of CCR5 inhibitors have been developed and display anti-HIV activity both *in vitro* and *in vivo*. These drugs are believed to work by binding to CCR5 at a site distinct from the gp120-binding site and subsequently alter the conformation of the CCR5 extracellular loops required for entry of R5-tropic HIV variants. One such CCR5 antagonist*, maraviroc,* is licensed for use in the United States and in Europe [43,72]. A variety of CXCR4 antagonists have also been developed, and while they exhibit potent anti-HIV activity *in vitro* (against X4 but not R5 virus strains)*,* administration of these drugs *in vivo* results in mobilization of HSCs from the bone marrow to the peripheral blood, highlighting the important role of CXCR4 in HSC homing [73-75]. Although this side effect limits their use in HIV-infected individuals, the CXCR4 antagonist *plerixafor*, is currently used to mobilize HSCs for subsequent autologous transfer in patients with non-Hodgkin's lymphoma and multiple myeloma [76].

gp41-mediated membrane fusion presents another drug target in the HIV entry process. Synthetic peptides based on the sequence of HR2 display significant antiviral activity against HIV *in vitro* but for many years, the mechanism of this antiviral activity remained unknown [77]. However, the

observation that these peptides display higher antiviral activity as dimers and the elucidation of the structure of the gp41 fusion machinery have led to a model for their mechanism of action [63,64,78]. These drugs are now believed to act in a dominant negative fashion by competing with the HR1 and HR2 domains of gp41 and ultimately preventing the formation of the six-helix bundle required for membrane fusion. One such peptide, *enfuvirtide,* is the only FDA-approved HIV fusion inhibitor, and is indicated for use in combination with standard antiretroviral therapy. However the twice-daily subcutaneous dosing schedule of the drug makes it an unattractive choice for many patients and care providers.

As is the case for most anti-HIV drugs, viral variants resistant to all of the entry inhibitors have been identified. The appearance of *maraviroc* resistant viruses *in vitro* is a well-established phenomenon and these viruses either adapt to recognize the drug-bound conformation of CCR5, or more commonly, acquire the ability to use CXCR4 in addition to CCR5 (reviewed in [79]). In vivo, resistance to *maraviroc* most commonly results from outgrowth of pre-existing X4 or R5X4 variants that are sometimes present at very low levels prior to the onset of therapy. As such, tropism testing is indicated prior to administration of *maraviroc*. In the case of *enfuvirtide,* mutations within HR1 and HR2 rapidly select for viruses with a dramatically decreased sensitivity to the drug [80,81]. HIV drug resistance is not unique to entry inhibitors, but is a widespread problem seen with all classes of HIV chemotherapeutics, and while ART increases survival in HIV-infected individuals, the problems of drug resistance, impaired

immune function despite ART, long-term financial cost and drug-associated toxicities of ART continue to fuel the search for curative therapies for HIV infection.

VII. HIV entry as a target for gene therapy

To date, several studies have explored genetic approaches to hijack or halt the 3 main steps of HIV entry—CD4 binding, coreceptor binding and membrane fusion. Outside of its role as the primary receptor for HIV, CD4 plays a critical role in antigen recognition by the T-cell receptor and as such, abrogation of CD4 expression is not a viable option to prevent HIV entry. A number of early gene therapy studies, however, took advantage of the fact that HIV-1 requires CD4 to enter cells, and coupled the extracellular and transmembrane domains of CD4 to the intracellular signaling domain of the invariant ζ-chain of the T cell receptor (TCR) thus pairing viral recognition by CD4 with TCR signaling and downstream effector functions. Introduction of these chimeric receptors into CD4 and CD8 T cells resulted in HIV-specific targeting by both cell types *in vitro*. In particular, expression of these chimeric TCRs in cytotoxic CD8 cells allowed them to recognize and kill HIV infected cells which often express Env on their surface (67,68). Following these promising preclinical studies, two clinical trials investigated the effects of adoptive transfer of chimeric TCR modified CD4 and CD8 T cells on HIV infection. In both trials, the gene-modified cells successfully engrafted and trafficked to the rectal mucosa—a

major site of HIV replication (69-71). While neither study observed a significant decrease in the viral load of treated subjects, one study reported a trend towards a decrease in reservoir size following treatment with the gene-modified cells(71).

To date, most gene therapy attempts to inhibit HIV entry have focused on interfering with the interaction between the virus and its coreceptors by either reducing or eliminating coreceptor expression. When attempting to genetically modulate coreceptor expression, an important consideration is the choice of cells to be treated as the ease with which target cells can be modified, and their longevity and capacity for self-renewal all influence the chances of success (reviewed in (72)). For this reason, most studies have focused on modifying either T cells or CD34+ HSCs with the eventual goal of adoptive transfer of the gene-modified cells. CCR5 is a particularly attractive target for HIV entry-focused gene therapy as the complete loss of CCR5 expression appears to be well tolerated in *ccr5*Δ32 homozygotes. Additionally, the CCR5 small molecule antagonist *maraviroc*, which is currently approved for use in HIV infected patients, exhibits potent antiviral activity without adversely affecting immune cell function, suggesting that a partial or complete loss of CCR5 may not result in severe immunologic consequences (64,65). On the other hand, less is known about the potential consequences of decreasing or completely ablating CXCR4 expression. In particular, CXCR4's role in the bone-marrow retention of HSCs (73) may result in the unintended side effect of HSC egress from the bone marrow into the peripheral blood.

A number of studies have demonstrated the ability of exogenous transgenes to target the HIV coreceptors at the level of protein, RNA or DNA, with the final common result being decreased surface coreceptor expression. Early work using a CCR5-specific single-chain antibody engineered to express an ER-retention motif showed that these “intrabodies” prevented trafficking of the CCR5 protein to the cell surface. When introduced into susceptible cells, the resulting intrabody-mediated intracellular sequestration of CCR5 resulted in a decrease in infection by CCR5-using HIV (74,75). RNA-based approaches have also provided promising results—multiple studies have used RNA interference, CCR5-targeted ribozymes or a combination of both to efficiently decrease levels of CCR5 mRNA and thus surface expression of CCR5 (76-82). A recent clinical trial examined the safety and potential efficacy of one such RNA-based agent by following the adoptive transfer of autologous CD34+ stem cells transduced with a lentivirus encoding a CCR5 ribozyme, an anti-HIV siRNA and an RNA decoy that prevents initiation of HIV transcription (83). In this study, the gene-marked cells engrafted successfully, and multiple hematopoietic lineages expressing the transgene were detectable for up to two years post-infusion. However, these cells did not provide any observable clinical benefit in terms of CD4 count or HIV viral load. While the lack of clinical benefit was likely in part due to the low percentage of gene-marked cells infused (~0.14% of infused cells were gene-marked on average), another potential problem with this and other RNA-based gene therapy approaches to decrease CCR5 expression is their inability to completely and permanently eliminate surface CCR5 expression. This poses a

problem, as many HIV-1 isolates are capable of using low surface levels of CCR5 to enter cells. For this reason, many groups have begun exploring permanent modification of the host genome using designer nucleases so as to completely eliminate surface CCR5 expression.

CCR5-specific ZFNs are capable of permanently inactivating the CCR5 gene in primary CD4 T cells, and ZFN-modification confers a survival advantage on gene-modified cells in the presence of CCR5-using HIV both *in vitro* and in a humanized mouse model of HIV infection (14). Additionally, ZFN modification of CCR5 in primary CD34+ HSCs results in the production of multiple hematopoietic lineages all lacking surface CCR5 expression (21). As a result of these promising preclinical data, several clinical trials are currently investigating the safety and efficacy of autologous transplants using ZFN-modified CD4 T cells or CD34+ HSCs in HIV infected individuals (www.clinicaltrials.gov; Trial identifiers NCT00842634, NCT01044654, NCT01543152). While these strategies are capable of generating cells that are highly resistant to infection with CCR5-using HIV, they offer no protection against viruses that use CXCR4. To this end, a recent study showed the ability of a ZFN pair targeting the CXCR4 coreceptor to specifically and efficiently inactivating this coreceptor in primary human CD4 T cells and in a humanized mouse model of HIV infection(15). However, in this study, protection of CD4 T cells in animals harboring CXCR4-modified cells was only transient, due to the outgrowth of viruses capable of using CCR5 to infect cells. These results, and the emergence of CXCR4-using HIV in patients who fail

treatment with the CCR5 antagonist *maraviroc,* highlight the importance of preventing both CCR5- and CXCR4-mediated entry in order to completely abolish infection by the two major HIV variants.

Chapter 3. Simultaneous zinc-finger nuclease editing of the HIV coreceptors *ccr5* and *cxcr4* protects CD4+ T cells from HIV-1 infection

Contributors to this work include Craig B. Wilen (UPenn), Jianbin Wang (Sangamo Biosciences), Jennifer Duong (Sangamo Biosciences), Anthony J Secreto (UPenn), Gwenn A. Danet-Desnoyers (UPenn), James L. Riley (UPenn), Phillip D Gregory (Sangamo Biosciences), Carl H. June (UPenn), and Michael C. Holmes (Sangamo Biosciences).

I. Statement of hypothesis

I hypothesize that the simultaneous inactivation of *ccr5* and *cxcr4* in T cell lines and primary human CD4 T cells is feasible. Additionally, I hypothesize that cells lacking both CCR5 and CXCR4 will have a survival advantage in the presence of both CCR5 and CXCR4-using HIV *in vitro.* Finally, I hypothesize that humanized mice bearing these double-gene modified cells will be better able to maintain their CD4 counts in the face of HIV-1 challenge as compared to wild-type mice, or mice lacking a single coreceptor.

II. Materials and methods

Cell culture and ZFN treatment. SupT1 T cells expressing multiple copies of CCR5 introduced by lentiviral transduction (SupT1-R5), and primary human

CD4+ T cells were maintained in RPMI growth media (Invitrogen) supplemented with 10% fetal bovine serum (FBS). CD4+ T cells were grown in the presence of human interleukin 2 (IL-2) (100 IU/ml). We used CCR5 (R5) and CXCR4 (X4) zinc finger proteins (ZFPs) from our previous studies(14,15). However, the ZFN nuclease domains were the ELD/KKR variant which were modified to function as obligate heterodimers with enhanced cleavage activity(84). To deliver ZFNs, we simultaneously transduced cells with two Ad5/F35 vectors encoding either the R5- or X4-ZFNs, or an Ad5/F35 GFP-expressing control vector(14,15). This is a chimeric adenoviral vector based on the AdEasy vector system (Promega) with an E1/E3 deleted backbone and a chimeric fiber gene comprised of a serotype 5 fiber tail domain, and serotype 35 shaft and knob domains(85). CD4+ T cells were activated 18-24 hours prior to vector transduction with anti-CD3/anti-CD28 coated magnetic beads(15) . We determined vector multiplicities of infection (MOIs) using the 293T cell line and measured mutation frequencies by surveyor nuclease (Cel1) or T7 Endonuclease I assay (T7E1), or deep sequencing(22,86,87).

Virus production and infection. Pseudoviruses mediate one round of infection as the viral glycoprotein gene is only supplied *in trans*. We generated HIV pseudoviruses encoding a GFP reporter, bearing either the HIV glycoprotein (Env) or the Vesicular Stomatitis Virus glycoprotein (VSV–G) by co-transfection of the viral glycoprotein and the pNL4.3-Δ*env*-vpr^{+}-eGFP-HIV backbone(88) into 293-T cells using calcium phosphate. We harvested supernatants after 72 hours

and concentrated virus by ultracentrifugation(89). The HIV pseudoviruses were made using the R5 and X4-HIV *envs* JRFL and TYBE(90) respectively. For SupT1-R5 pseudovirus infections, we infected $1x10^5$ cells in a 96-well v-bottom plate with 500ng of HIV-1 p24 or 5-10ng of HIV-VSV-G p24. Samples were spun at 1200x*g* for 90 minutes at 25°C then transferred to 37°C. Infection was measured at 96-hours by flow cytometric analysis of GFP positive cells.

Replication competent HIV-1 was made in primary CD4+ T cells (89) and we infected SupT1-R5 cells and primary cells with 50-100ng p24 per million cells 5-7 days following ZFN treatment. The replication competent HIV-1 strains used were the R5-virus BaL, and the X4-viruses BK132 and HxB2. All CD4+ T cell infections were performed using cells from three independent donors. For all infections where cells were challenged simultaneously with R5 and X4-HIV, the viruses were mixed in a 1:1 ratio normalized by HIV-1 p24.

Flow cytometry. Cells were stained at room temperature in Fluorescence-activated cell sorting (FACS) buffer (PBS-/-, 2% FBS, 2mM EDTA). SupT1-R5 cells and CD4 T cells were stained with Live/Dead Aqua and QD655-CD3 (clone S4.1) (Invitrogen), Fluorescein isothiocyanate(FITC)-CD45 (clone H130) (Biolegend), AlexaFluor700-CD4 (clone RPA-T4), PacificBlue-CD8 (clone RPA-T8), Allophycocyanin (APC)-CCR5 (clone 2D7) and Phycoerythrin (PE)-CXCR4 (clone 12G5) (BD Biosciences). Samples were run on an LSR II (BD Biosciences), and analyzed using FlowJo 10.0.4 (Treestar).

HIV-1 humanized mouse challenge.

12-14 week old NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/Szj (NSG) immunodeficient mice received either unmodified CD4 T-cells, R5-ZFN-treated CD4 T-cells, or R5/X4-ZFN-treated CD4 T-cells. Each group contained 18 animals and each animal received $1.8x10^7$ cells. Mice were randomly assigned to control for cage, sex and birth cohort effects. The animals were maintained in a defined flora facility at the University of Pennsylvania, with approval of the Institutional Animal Care and Use Committee. We measured peripheral blood CD4 counts 21 days post-infusion to assess engraftment by staining for CD45+/CD3+/CD4+/CD8- cells. Three days later, half the animals in each group received $5x10^4$ CD4+ T-cells infected with the R5-virus BaL, and $5.0x10^4$ CD4+ T-cells infected with the X4-virus Bk132. Control animals received $1x10^5$ uninfected CD4+ T-cells. We performed retro-orbital bleeds to assess CD4 counts and measure gene disruption. Mice were euthanized following the development of xenogeneic-graft-versus-host-disease. Cardiac punctures and splenectomies were performed on all mice following euthanasia. We passed spleens through a 70μm strainer to obtain a single-cell suspension for measuring gene modification and CD4 count. ZFN-induced mutation frequencies were determined by performing deep sequencing at the R5 and X4-ZFN target sites(22).

Off-target site analysis. We previously used systematic evolution of ligands by exponential enrichment (SELEX) to determine the *in vitro* binding preference of

each zinc finger protein (ZFP)(14,15). Using the information on CCR5 and CXCR4 binding preferences, we generated a position-weight matrix which we aligned to the human genome to look for sites with similarity to each of the four possible combinations of CCR5-CXCR4 cross heterodimers (R5-left with X4-right; R5-left with X4-left; R5-right with X4-right; R5-right with X4-left). We allowed up to a 4bp mismatch compared to the SELEX consensus sequence, and considered sites with a 5bp or 6bp spacer between each ZFP pair. We performed deep sequencing at these sites to identify off-target activity(22). Off-target activity was defined as an insertion/deletion (indel) involving the target region (20 bp) centered at the nuclease binding sites at a frequency of greater than 0.1%, and with a calculated ratio of cleavage (Z-M)/M of greater than 2, where Z represents % indels in ZFN treated samples, and M represents % indels in mock treated samples.

III. Results

Simultaneous ZFN disruption of *ccr5* and *cxcr4* in a T-cell line protects from R5 and X4-HIV infection

To determine whether two ZFN pairs targeting the HIV-coreceptors *ccr5* (R5) and *cxcr4* (X4) could abolish expression of both coreceptors, we co-transduced the human SupT1-R5 T-cell line with increasing amounts of the Ad5/F35 vectors encoding the R5-ZFN and X4-ZFN. Co-administration of both (R5/X4) ZFNs caused a dose-dependent reduction in cell surface expression of

both coreceptors with 9% of cells no longer expressing either coreceptor at the highest MOI, while delivery of the same vector encoding GFP had no effect on surface R5 or X4 levels (Figures 3A-B). Next, we assessed the ability of *ccr5* and *cxcr4* co-disruption to protect cells from infection with R5 and X4-using HIV-1 strains. We challenged untreated or R5/X4-ZFN treated SupT1-R5 cells simultaneously with the R5- and X4-HIV strains BaL and BK132 respectively and monitored surface coreceptor expression for 42 days. The proportion of double (R5/X4) negative cells increased over the course of the infection such that by 25 days post-infection, 96–99% of the R5/X4-ZFN treated cells challenged with HIV no longer expressed either coreceptor (Figure 3C-D). In contrast, the proportion of double-negative cells in the uninfected R5/X4-ZFN treated group did not change significantly over this time (Figure 3C, top panels). We also monitored cell viability and while untreated cells all died by ten days post-infection, a portion of the R5/X4-ZFN treated cells survived the infection with cell numbers increasing exponentially over time (Figure 3E). The coreceptor negative cells that survived HIV infection were re-challenged with pseudovirus bearing either HIV Env or VSV-G to confirm that the mechanism of resistance was due to inhibition of HIV Env-specific entry. These cells were highly resistant to either R5 or X4-HIV pseudovirus re-challenge, displaying 170-fold and 92-fold decreases in reinfection with R5- and X4-HIV respectively (Figure 3F). Of note, low levels of residual infection (~0.1%) were observed in the R5/X4-ZFN group pre-challenged with HIV. However, the majority of this residual infection occurred in the 1-4% of cells still expressing CCR5 and CXCR4 (data not shown). Additionally, these

coreceptor negative cells were readily infected by HIV expressing the vesicular stomatitis virus glycoprotein (VSV-G HIV), a virus that is not dependent upon R5 or X4 for infection (Figure 3F). Taken together, these data suggest that simultaneous delivery of two ZFN pairs is an efficient and viable strategy to disrupt *ccr5* and *cxcr4,* resulting in cells resistant to infection with both R5 and X4-using HIV.

ZFNs achieve simultaneous disruption of *ccr5* and *cxcr4* in human CD4+ T cells

To determine the feasibility of this approach in primary cells, CD4+ T cells from healthy human donors were simultaneously transduced with the adenoviral vectors encoding the R5 and X4-ZFNs. Transducing cells with MOIs of up to 500 of each ZFN resulted in modification of approximately 20% of all R5 and X4 genes as measured by the Cel1 assay, without significant impact on cell growth compared to untransduced controls (Figure 4A-B). As surface expression of CCR5 is undetectable on most primary CD4+ T cells, we could not determine the proportion of cells lacking both coreceptors by flow cytometry. To exclude the possibility that co-administration of both ZFNs in primary cells results in R5 and X4 gene modification in mutually exclusive cells, we delivered both ZFNs to CD4+ T cells and sorted them by FACS based on surface CXCR4 expression into CXCR4-high and CXCR4-low populations (Figure 4C). We then measured

levels of R5 and X4 gene modification in both of these populations. While *cxcr4* gene modification was greatly reduced in the X4-high population following sorting, levels of *ccr5* modification were similar in the X4-high and X4-low populations (Figure 4D), suggesting that adenoviral dual-ZFN treatment achieved dual-modification of *ccr5* and *cxcr4* in primary CD4+ T cells.

Depletion of central memory CD4+ T cells is a hallmark of HIV pathogenesis and progression to AIDS(91). As such, protection of this long-lived subset is important for the success of any HIV gene therapy approach. We therefore determined whether central memory CD4+ T cells could be modified by the R5- and X4-ZFNs. We first transduced CD4+ T cells with both ZFNs, expanded them *in vitro* for 10 days and then sorted the cells by FACS into central and effector memory populations based on surface expression of the memory markers CCR7 and CD45RO(92) (Figure 4E). CD4+ T cell activation using CD3/CD28 coated beads results in differentiation of ~97% of cells into a memory phenotype, with roughly half of all cells being either central or effector memory cells and the proportion of effector and central memory subtypes is not affected by treatment with the adenoviral vector or the ZFN (data not shown). We detected both *ccr5* and *cxcr4* modification (Cel1) in the central memory (CCR7+, CD45RO+) and effector memory (CCR7-, CD45RO+) populations at similar levels (Figure 4E) suggesting that we can effectively target the coreceptors in these critically important T cell subsets.

Survival advantage of dual-ZFN treated CD4+ T cells *in vitro*

To determine if the level of gene modification seen in primary cells rendered them resistant to both forms of HIV-1, we treated CD4+ T cells with both ZFNs, the R5-ZFN alone, or with a GFP control. We then challenged all three groups with a mix of R5- and X4-using HIV-1 and monitored cell growth and viability. By 32 days post-infection, there were no detectable live cells in the groups that received either no ZFN or the R5-ZFN alone, whereas cells from the R5/X4 ZFN group continued to expand (Figure 5A). We measured levels of *ccr5* and *cxcr4* gene modification 3 weeks into the infection in the R5/X4-ZFN treated group and observed increases in both *ccr5* (1.8-fold) and *cxcr4* (1.9-fold) gene modification (Figure 5B-C) when compared to the start of the infection. The preservation of cell growth and increases in gene modification seen only in the R5/X4-ZFN group in the presence of HIV suggest that ZFN-modification of both coreceptors provides a significant survival advantage in the presence of R5- and X4-using HIV.

Analysis of off-target cleavage following dual ZFN administration

Simultaneous administration of two ZFN pairs creates the theoretical possibility of forming four different ZFN cross-heterodimers made up of the left and right halves of the two unique ZFN pairs. These ZFN cross-heterodimers may subsequently bind to unintended target sites, resulting in more off-target

gene modification compared to delivery of a single ZFN pair. As the target sequence contributes to ZFN binding and thus specificity(93), we previously determined the *in vitro* DNA binding preference of each CCR5 and CXCR4 ZFP using SELEX(14,15). We constructed a position-weight matrix for each 12bp ZFP binding site, and by comparing these sequences against the human genome to identify potential off-target binding sites, identified sites that were in fact independently cleaved by the R5- and X4-ZFNs(14,15). In this study, we took a similar approach to identify off target sites that could result from the binding and cleavage of a ZFN heterodimer consisting of one half R5-ZFN and one-half X4-ZFN. We performed deep sequencing at these predicted sites in primary CD4+ T cells following treatment with both the R5- and X4-ZFNs compared to otherwise identical untransduced controls (Table 2). Using genomic DNA from samples where on-target modification of *ccr5* and *cxcr4* were 28% and 20% respectively, we analyzed 4,000-20,000 reads per predicted off-target site but failed to detect significant levels of modification at any of the 40 predicted off-target sites (Table 2). These results suggests that cross-heterodimerization of the R5- and X4-ZFN pairs following delivery to primary CD4+ T cells does not result in detectable off-target activity, at the 40 sites examined. Moreover, the use of recently described orthogonal heterodimeric nuclease domains that further restrict cleavage (94) could further address this concern.

Dual coreceptor disruption protects CD4+ T cells following HIV infection *in vivo*

NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/Szj (NSG) immunodeficient mice are a useful model for studying HIV-1 *in vivo* as their immune systems can be reconstituted using human cells that are susceptible to HIV infection(95). To assess whether our approach protects against infection with both R5- and X4-HIV *in vivo*, we treated primary CD4+ T cells with no ZFN, the R5-ZFN alone or with the R5- and X4-ZFNs simultaneously. Three groups of 18 mice each were infused with $1.8X10^7$ cells per animal from each treatment group (Figure 6A). We observed CD4+ T cell engraftment across all groups 21 days post-infusion (Figure 6B). We then infected half the animals in each group by simultaneously infusing $5x10^4$ unmodified (no ZFN) CD4+ T cells infected with the R5-virus BaL and $5x10^4$ unmodified CD4+ T cells infected with the X4-virus BK132. Of note, while the BK132 HIV-1 swarm primarily uses CXCR4 to enter cells, it can also achieve very low levels of infection using CCR5(96) as evidenced by the outgrowth of R5-HIV when this virus was used to challenge X4-ZFN treated CD4 T cells(15). Prior to infusion, the frequency of HIV p24+ cells was 58% and 44% in the R5 and X4-HIV infected cells respectively. Control animals in each group received an infusion of $1x10^5$ uninfected, unmodified CD4+ T cells. Animals that received R5-ZFN or R5/X4-ZFN treated cells had 5.5 and 12.8-fold lower CD4 counts respectively, than animals that received untransduced cells when measured three days prior to infection (Figure 6B). However, this difference in CD4 counts

was no longer present at the time of the first peripheral blood sampling six days post-infection (Figure 6D). This slight engraftment difference was not seen previously (14,15) and may be donor specific. While CCR5 ablation alone was previously shown to protect against challenge with CCR5-using HIV(14), animals that received both ZFNs were better able to maintain their CD4+ T cell counts following challenge with R5- and X4-HIV (Figure 6D). Specifically, CD4 counts in animals from the infected R5/X4-ZFN group were 35 and100-fold higher than those in animals from the mock and R5-ZFN infected groups respectively by 22 days post-infection, and these differences increased to an average 200-fold difference 55 days post infection (Figure 6D). Additionally, we measured the proportion of human CD4+ T cells in the spleens of infected animals at time of sacrifice. In animals that received dual-ZFN treated cells, up to 10% of all cells populating the spleen were human CD4+ T cells whereas we could not detect any human CD4+ T cells in spleens from animals that received either no ZFN or the R5-ZFN alone (Figure 6C).

To determine the mechanism of CD4 count maintenance observed in the R5/X4-ZFN group, we performed deep sequencing at the ZFN cut-sites in CD4+ T cells from animals in the R5/X4-ZFN group to measure changes in the proportion of ZFN-induced indels. There was no change in *ccr5* or *cxcr4* gene modification in uninfected animals that received both ZFNs. However, on average, 69% of all *ccr5* genes and 73% of all *cxcr4* genes were mutated in CD4+ T cells from the infected R5/X4-ZFN animals by 34 days post-infection

(Figure 7A). This represents a 2.2-fold increase in *ccr5*-modification and a 2.3-fold increase in *cxcr4*-modification compared to six days post-infection. Additionally, we observed significantly lower surface expression of both coreceptors on cells from the infected R5/X4-ZFN treated mice 34 days post-infection (Figure 7B). These data provide strong evidence that simultaneous disruption of *ccr5* and *cxcr4* using ZFNs is capable of generating a pool of cells that are resistant to the major forms of HIV *in vivo*, and these cells can engraft, traffic normally to the spleen, and have a significant survival advantage in the presence of R5- and X4-using HIV.

IV. Discussion

Genetic approaches to control HIV infection have so far involved the removal of host genes required by the virus, or the introduction of antiviral genes that interfere with virus replication(72,97). While both are attractive options, the latter approach could be hampered by the potential immunogenicity of foreign transgenes and the diversity of the HIV viral quasispecies in chronically infected individuals, which may allow for the outgrowth of viral variants resistant to these antiviral genes. Removal of essential host factors such as the coreceptors, may present a greater challenge for the virus to overcome as suggested by the natural HIV resistance in *ccr5*Δ32 homozygotes and the functional cure achieved following the transplantation of *ccr5*Δ32 stem cells(59,60). Genetic editing through the use of coreceptor-specific ZFNs clearly renders a fraction of cells

coreceptor-negative, as HIV-1 challenge of ZFN-treated cells results in preferential survival of edited cells lacking the targeted coreceptor(14,15). Treatment with R5-specific ZFNs results in permanent modification of 30% and 50% of *ccr5* alleles in human hematopoietic stem cells and T cells respectively(14,22), with a bi-allelic modification frequency of up to half of that(22). The efficiency of bi-allelic disruption becomes even more important when attempting to disrupt both *ccr5* and *cxcr4*, necessitating the inactivation of four genes in a given cell to render it resistant to infection by virtually all HIV-1 strains.

Individuals lacking *ccr5* do not display severe phenotypic abnormalities and its pharmacologic blockade is well tolerated(64), making it an ideal target for permanent modification. However less is known about the effects of the specific loss of CXCR4 expression on CD4+ T cells. One study has tried to address this by generating a T cell-specific knockout of *cxcr4* in mice(98). These animals were born in normal numbers and displayed normal T cell function compared to wild-type mice, but displayed defective responses to the CXCR4 ligand, SDF-1. We have previously shown that our CXCR4-ZFN is capable of efficiently modifying CXCR4 in rhesus macaque cells, providing us with the unique opportunity of examining the engraftment and trafficking of CXCR4-ZFN edited CD4+ T cells in a more relevant model system before attempting this approach in humans(15).

One potential limitation of this study is the modification of only CD4+ T cells and not a longer-lived cell population such as hematopoietic stem cells (HSCs). While previous studies have shown that *ccr5* can be efficiently modified in CD34+ HSCs and this modification results in the production of *ccr5* null CD4+ T cells and macrophages that are resistant to infection, CXCR4 signaling plays a critical role in the homing of HSCs to the bone marrow and as such, loss of CXCR4 expression on HSCs may result in their unwanted egress into the peripheral blood(73). However, due to the long half-life of memory CD4+ T cells, we believe that an approach consisting of modification of *ccr5* in HSCs and both coreceptors in CD4+ T cells will provide patients with a pool of cells resistant to R5-HIV infection while at the same time preventing the outgrowth of viruses capable of using CXCR4.

V. Acknowledgements

I thank Beatrice Hahn and Nicholas Parrish for providing HIV-1 virus stocks, and James Hoxie for generously providing the SupT1-R5 cell line. I also thank the University of Pennsylvania's Center for Aids Research (CFAR) supported Human Immunology core for human CD4+ T cells, and the Stem Cell and Xenograft core. Additionally, I thank Xiaolong Fong for generously providing the Ad5/F35 vector, Frederick Bushman, Jason Wojcechowskyj, Nicholas Parrish, and Zahra Parker for helpful comments and technical advice. I also thank David Shivak, Kenneth Kim, and Jenny Jiacheng Yan at Sangamo Biosciences Inc. for help in analyzing on- and off-target gene modification in ZFN-treated samples.

CHAPTER 4. Translocation capture sequencing reveals the genome-wide specificity of ZFNs, TALENs and CRISPR/Cas associated nucleases

Contributors to this work include Mary J. Drake (University of Pennsylvania), Nirav Malani (University of Pennsylvania), Craig B. Wilen (University of Pennsylvania) and Frederic Bushman (University of Pennsylvania)

I. Introduction

Precision genome engineering allows for the introduction of stable, heritable genetic modifications into complex organisms in a sequence specific manner. These modifications are accomplished using one of the three main classes of designer nucleases—zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeat (CRISPR) associated nucleases. Designer nucleases hold great promise for the treatment of human disease, but a high level of nuclease specificity is likely required to ensure that their use does not result in genotoxic effects. To this end, a number of studies have sought to identify sites of unwanted or off-target cleavage activity for each of these nucleases (14,15,25,99-103). However, most of these studies either employ approaches predicated on the assumption that sequence similarity is the primary driver of off-target activity, or attempt to predict sites of off-target activity based on data obtained from *in silico* systems that may not faithfully recapitulate the DNA binding events that occur in a living cell. While all of these studies have identified sites of off-target activity with sequence similarity to the target site, a recent study

took a slightly different approach by co-administering a ZFN and an integration defective lentiviral vector, and showed that this vector was able to capture sites of ZFN-induced double strand breaks (DSB)(26). Using this approach the authors showed that ZFNs are able to cleave sequences with as low as 67% homology to their on-target sites and importantly, many sites with close to 100% homology to the ZFN target site showed no signs of nuclease activity (26). These results suggest that sequence specificity may not be the sole determinant of nuclease-induced DNA double strand breaks (DSBs). We thus sought to unambiguously identify all nuclease induced DSBs in cells treated with each of the three main classes of designer nuclease using a more unbiased approach. We hypothesized that the simultaneous introduction of two DSBs within a cell by a designer nuclease—at on and off-target sites—may result in the formation of translocations between these sites, and we attempted to identify these translocations using Translocation Capture sequencing (TCseq). TCseq allows the high throughput identification of chromosomal rearrangements across the genome (104-107). Using this technique, a recent study showed that translocations in B-lymphocytes are associated with the rate of repeated site-directed DNA damage (104-107). In this study, we set out to ask the following questions: (1) Do double strand breaks generated by a single designer nuclease at on and off-target loci result in the formation of translocations between these breaks? (2) Can these translocations be detected using TCseq? (3) Will the genome-wide characterization of the off-target sites involved in these

translocations further our understanding of the factors governing the specificity of these nucleases?

II. Materials and methods

Designer nucleases. In this study, we employed a CCR5-specific Compozr® Zinc Finger Nuclease (Sigma Aldrich) and a CCR5-specific XTN-TALEN (Transposagen Biopharmaceuticals). Both the ZFNs and TALENs encoded the ELD:KKR FokI nuclease domains with obligate heterodimeric FokI activity(84). CRISPR guide RNAs (gRNA) targeting the *ccr5, cxcr4* and *vegfa* genes were designed based on the architecture described in (12), and the *vegfa* sequences targeted were previously described (103). We obtained a human codon-optimized *S. pyogenes* Cas9 encoding plasmid from Addgene (Plasmid 41815). The Cas9 coding sequence was excised from pCDNA3.3 using XbaI and AgeI and blunt cloned into the BamHI site in the pGEM.64A vector. mRNA Transcripts derived from the resulting vector (termed pGEM.64A.hCas9) encode a 64-nucleotide long polyA tail(108). Sequences targeted by the ZFNs, TALENs and gRNAs are as outlined in Table 3.

***In vitro* transcription.** To generate GFP, ZFN, TALEN, and Cas9 mRNA's, plasmids encoding the respective mRNAs were linearized (XbaI for the ZFN left and right plasmids, PmeI for the TALEN left and right plasmids, SpeI for the pGEM.64A.hCas9 and pGEM.GFP.64A plasmids), and transcribed using the MEGAscript T7 *in vitro* transcription kit (Life technologies). Reactions were

performed at 37^{o}C for 16 hours as per manufacturer's instructions. To ensure proper capping of transcripts, transcriptions were performed using the 3´-0-Me-m^{7}G(5')ppp(5')G RNA Cap Structure Analog (New England Biolabs) at a 4:1 cap analog to GTP ratio. ZFN and TALEN plasmids were subsequently polyA-tailed using the A-plus Poly(A) Tailing Polymerase kit (Cellscript). To synthesize a gRNA mRNA targeting *cxcr4,* polymerase chain reaction (PCR) was used to simultaneously amplify the 100bp *cxcr4* gRNA and add a T7 promoter to the 5'-end of the PCR amplicon (Table 5). This amplicon was subsequently used as a template for T7 *in vitro* transcription without a cap analog, as described above. All mRNA transcripts were purified by lithium chloride precipitation.

Cell culture and designer nuclease delivery. 293T cells were maintained in DMEM growth media with 10% fetal bovine serum (FBS). For each *vegfa* CRISPR transfection, triplicate wells of 4x10^{5} 293-T cells were plated in a 12-well plate. After 24 hours, these cells were transfected with either a GFP encoding plasmid (pHyg-EGFP, Clontech), or with 5ug each of two plasmids encoding the desired gRNA (pCRII Blunt TOPO) and the human codon-optimized *S. pyogenes* Cas9 (in pCDNA3.3) as previously described (12). Transfections were performed using Lipofectamine 2000 (Life Technologies) as per manufacturer's instructions. Cells were transferred to a 6-well plate 24 hours post-transfection and were harvested 5 days post-transfection. Each condition was transfected in triplicate in each of two independent experiments. Genomic DNA from independent

experiments (~30 million cells per condition) was pooled for sequencing library preparation.

Jurkat T cells (clone E6-1) were maintained in RPMI growth media (Invitrogen) supplemented with 10% FBS. ZFNs, TALENs, and CRISPR/Cas nucleases were delivered to Jurkat T cells by electroporation using a BTX-ECM830 electroporator (Harvard apparatus BTX, Holliston, MA). Jurkat T cells were washed three times in OPTI-MEM (Invitrogen) and resuspended at a concentration of $1x10^8$ cells/mL in OPTI-MEM. For each condition, ~$1x10^7$ cells were mixed with varying amounts of the corresponding nuclease mRNA or DNA (Table 5) and electroporated in a 0.2cm electroporation cuvette (500V for 700μs). Cells were immediately transferred to a 6-well plate containing pre-warmed RPMI+10%FBS at a concentration of $2x10^6$ cells/mL and incubated at 30^{o}C for 24 hours. Following this, cells were maintained at a concentration of $1x10^6$ cells/mL for an additional nine days in a 37^{o}C incubator prior to genomic DNA extraction for gene modification analysis and sequencing library preparation. For each condition, genomic DNA from 150-250 million cells obtained from three independent experiments was pooled for sequencing library preparation.

Quantification of gene modification. Genomic DNA extractions were performed using the QIAamp DNA Blood Maxi kit (Qiagen) as per the manufacturers instructions. Mutation frequencies were measured as percent non-homologous end joining (%NHEJ) using the T7 Endonuclease I assay (T7E1) as

previously described (87). Primers used to perform the T7E1 assay are as outlined in Table 4.

Translocation Capture sequencing

Sequencing library preparation. For each sample, three 5µg aliquots of isolated genomic DNA were fragmented by enzymatic digestion using NEBNext dsDNA Fragmentase (New England Biolabs) to obtain a size distribution of 100-1000bp. Fragmented genomic DNA was purified using the QIAquick PCR purification kit (Qiagen), and purified DNA was subsequently blunted using the Quick Blunting™ Kit (New England Biolabs), and adenosine-tailed using the klenow 3' → 5' exonuclease (New England Biolabs). Each tailed fragment was ligated to a unique partially double-stranded DNA linker using a Quick ligation kit (New England Biolabs) to reduce sample contamination as previously described(109). Linker-ligated samples were pooled and amplified using Phusion DNA polymerase (New England Biolabs). We performed a first round PCR using a gene-specific primer located upstream of the nuclease induced double strand break, and a linker specific primer. A second PCR was performed using 1µl of a 1/50 dilution of the first round PCR as a starting template. The second round PCR primers were internal to those used in the first round PCR and contained barcodes to allow for sample multiplexing, and sequences necessary for binding to an illumina sequencing flow cell as previously described (110). Library amplification primers are as outlined in Table 6. For Jurkat T cell samples, 8 unique first round PCRs were performed and each of these was used as a

template for the second round PCR. For each 293T samples, two first round PCRs were performed and each of these was used as a template for four second round PCRs. All second round PCR samples were purified using the Agencourt Ampure XP beads (Beckman Coulter) and equimolar amounts of each sample were pooled to obtain a final sequencing library.

Data analysis. Paired ends were rejoined and the sequence end originating within the nuclease-treated gene was aligned to the 30bp genomic sequence immediately following the sequencing primer, allowing for up to a 2bp mismatch. This sequence was then trimmed and the remaining sequence was aligned to the human genome using BLAST with a 90% identity criteria. Two or more reads were necessary to score a putative translocation event and translocation frequency was determined by analyzing the number of unique fragment lengths as previously described(109).

Rearrangement verification. Nested PCR was used for both *de novo* identification and for verification of genomic rearrangements. Forward primers were always designed to anneal to the nuclease target gene upstream of the nuclease cut-site, while reverse primers were designed to bind the suspected translocation partner (Table 8). Second round PCR products were cloned into the pCR Blunt II TOPO vector using the Zero Blunt PCR TOPO kit (Invitrogen). Following transformation of cloned products individual colonies were analyzed by Sanger sequencing using the T7-primer included in the kit.

III. Results

Activity of CCR5-specific ZFNs, TALENs and CRISPRs with an overlapping target sequence

To enable us compare the specificity of the three main classes of designer nucleases, we designed and obtained a ZFN, TALEN and CRISPR/Cas nuclease targeting an overlapping sequence in CCR5 (Figure 8A). Delivery of increasing amounts of these nucleases to the Jurkat T cell line resulted in a dose-dependent increase in the level of gene modification at the target loci (Figure 8B). Additionally, we simultaneously treated these cells with two CRISPRs—one targeting *ccr5* and one targeting *cxcr4*—to determine whether simultaneous delivery of two designer nucleases resulted in an exponential increase in the rate of off-target cleavage, and to serve as a translocation control for future experiments. Using nested PCR and Sanger sequencing, we were able to identify reciprocal translocations between the *ccr5* and *cxcr4* loci involving the nuclease-induced double stranded breaks at these sites (Figure 9A-B)

A single CRISPR/Cas nuclease generates rearrangements between sites of on and off-target activity

Several recent studies have identified sites of CRISPR/Cas nuclease off-target activity(99,101-103). As such, we asked whether we could identify

rearrangements between these known on- and off-target loci. We employed two recently described CRISPRs targeting two unique loci in *vegfa* (termed VEGFA.T2 and VEGFA.T3). Delivery of either one of these nucleases to 293T cells results in gene modification at their respective target loci and at several off-target loci (103) (Figure 10A). Thus we after delivering either nuclease to 293T cells, we used nested PCR in an attempt to identify rearrangements between these loci and their previously identified off-target sites. We assayed for 2 rearrangements per CRISPR/Cas nuclease and were able to detect rearrangements of the expected sizes between all four on and off-target loci tested (Figure 10B). Using sanger sequencing, we confirmed that these rearrangements did in fact involve the expected loci (Figure 11A-D).

IV. Discussion

In this ongoing study, we attempted to identify translocations involving designer nuclease-induced double strand breaks throughout the genome, in the hope that identification of translocation partners would reveal off-target double strand breaks caused by these nucleases. We have currently completed our deep sequencing of the nuclease-treated and control samples and are currently completing the analysis of this sequence data. However, our interim results prove that delivery of a single CRISPR/Cas nuclease results in the formation of detectable rearrangements between sites of on and off-target gene modification. As such, we expect that the results of our TCseq data analysis will be useful in

the identification of sites of designer-nuclease induced off-target activity across the genome.

CHAPTER 5. Conclusion and Future directions

The focus of this thesis was two-fold—to explore the feasibility of simultaneously inactivating the two entry factors necessary for infection by the major forms of HIV using one class of designer nucleases (ZFNs), and to develop a tool for the genome-wide identification of translocations caused by the major classes of designer nucleases in an attempt to better understand the factors influencing the specificity of these nucleases.

In the first project, we showed that simultaneous delivery of two ZFNs targeting the HIV entry coreceptors *ccr5* and *cxcr4* results in stable inactivation of these genes in primary human CD4 T cells—the primary target of HIV infection. We then went on to show that these gene-modified cells are protected from infection by viruses that use either CCR5 or CXCR4 to infect cells, and have a survival advantage in the presence of these viruses. Using a humanized mouse model, we showed that animals harboring cells that lacked both coreceptors were better able to maintain their CD4 counts following infection with viruses that use both coreceptors. In fact, following virus challenge, animals with the gene modified cells had greater than 100-fold higher CD4 T cell counts on average than animals with wild-type cells. Additionally, analysis of the gene-modified cells following *in vivo* challenge revealed enrichment for gene-modified cells only in the presence of infection—supporting the hypothesis that knockout of both of these entry factors confers a survival advantage on primary CD4 T cells lacking these genes in the face of HIV infection. These results have

important implications for the field of HIV cure research. Specifically, our study has attempted to both phenocopy the remarkable cure achieved following the allogeneic transfer of stem cells from a *ccr5*Δ32 homozygous donor to a wild-type HIV-infected recipient(59,111) and extend upon those results by generating cells resistant to HIV-1 viruses that use either coreceptor (CCR5 or CXCR4) to infect cells. In this study, we utilized a humanized mouse model, which is a useful one for studying HIV infection *in vivo.* However*,* we will conduct future studies in a rhesus macaque infection model that more accurately recapitulates human HIV infection. The macaque model will allow for a more thorough investigation of the effects of our approach on HIV infection, including its effects on viral load, the DNA viral reservoir, and its ability to reconstitute the immune system—all in the setting of a prolonged antiretroviral therapy treatment interruption. In this model, we will also be better able to test the effect of loss of these coreceptors on T cell functionality by measuring the engraftment, trafficking, and longevity of these cells over an extended period of time *in vivo*. This model will also allow us to initially infect the animals, and subsequently perform an autologous transplant of gene-modified cells, to more closely mimic how this therapy will be administered in humans. Finally, this model will allow us test our long-term approach of modification of *ccr5* in both CD34+ hematopoietic stem cells and CD4 T cells, and modification of *cxcr4* in CD4 T cells alone.

Although inactivation of the viral coreceptors appears to be a logical approach to control HIV infection, what is less obvious is the amount of

coreceptor ablation required to observe a therapeutic effect. Data from *ccr5*Δ32 heterozygotes suggests that even partial coreceptor ablation may be of clinical benefit as these individuals display delayed disease progression following HIV infection(112). As such, the current ZFN technology may provide a therapeutic benefit despite not completely eliminating coreceptor expression. In support of this, in a recent phase I study of HIV-infected patients who received autologous R5-ZFN modified CD4+ T cells, a number of these patients exhibited viral load decreases compared with pretreatment levels following a 12-week interruption of ART (Carl H. June, University of Pennsylvania, written communication, May 23, 2013). Of note, the viral load in one of these patients reached undetectable levels prior to the reinstatement of ART and further analysis revealed this patient to be a *ccr5*Δ32 heterozygote. Despite the modest number of patients enrolled in this study, it is tempting to speculate that even the relatively low levels of coreceptor ablation achieved by ZFN modification may be of therapeutic benefit, and that this benefit will only increase as the expression of either coreceptor is further reduced.

A potential limitation associated with any approach that targets CCR5 alone is the presence of virus strains that utilize CXCR4. Clinical failure associated with the use of the CCR5 antagonist maraviroc is either the result of an outgrowth of pre-existing virus strains that use CXCR4 and continue to replicate despite CCR5 blockade(65), or mutations that arise in the viral glycoprotein allowing it to use the drug-bound form of CCR5(113), a resistance

pathway not available to the virus when *ccr5* is disrupted by ZFN technology. The fact that X4 HIV-1 strains can be present at low levels in chronically infected individuals provides a powerful rationale for employing a strategy that targets both *ccr5* and *cxcr4*(114). While R5 ZFNs have been used to ablate *ccr5* in hematopoietic stem cells (HSCs) to provide a self-renewing population of CCR5 negative cells, CXCR4 signaling plays a critical role in the bone marrow retention of HSCs(115,116). As a result, loss of CXCR4 expression on HSCs may result in their unwanted egress into the peripheral blood. Although less is known about the effects of the specific loss of CXCR4 expression on CD4 T cells, a recent study of T cell-specific *cxcr4* knockout mice showed that these animals are born in normal numbers and display humoral and cellular responses indistinguishable from those of wild-type mice(98), suggesting that loss of CXCR4 expression on CD4 T cells may be immunologically tolerated.

Our work and that of others provides compelling evidence for the ability of ZFN-mediated coreceptor ablation to protect CD4+ T cells and provide virologic control of HIV infections. However, a number of questions remain regarding the clinical efficacy of this approach and the ease with which it can be implemented. While the original transplant using *ccr5*Δ32 cells resulted in a cure, other factors may have contributed to this outcome including reductions in reservoir size by the conditioning chemotherapy and graft-versus-host disease experienced by the patient. Allogeneic HSCT has long been known to exert an anti-tumor or graft vs tumor effect (117). While there is conflicting evidence on the effect of allogeneic

HSCT on HIV infection (reviewed in (118,119)), the idea that complete donor chimerism, if achieved following an allogeneic HSCT, would effectively cure an established HIV infection is an interesting one. For this to occur, the donor cells would likely need to be protected from infection during the transition to complete donor chimerism, and in the case of the *ccr5*Δ32 transplant that resulted in a cure, that was achieved by using cells resistant to HIV infection. Another instance where protection of newly infused donor cells may have provided a clinical benefit in HIV infection is in the recent case of the 'Boston patients.' In this small study, three HIV-positive patients in Boston received allogeneic HSCT following reduced-intensity chemotherapy to treat lymphoma. The patients were kept on ART for up to 4 years following transplant and while only two patients survived, they were both shown to have a significantly lower viral reservoir as measured by viral outgrowth following maximum *in vitro* T cell activation (66). Moreover, these patients have recently undergone ART interruption and while results are only available for a few months post-ART interruption, there is no evidence of viral rebound as measured by viral RNA or DNA in the peripheral blood and rectal tissues of either patient [91]. This report is remarkable for two reasons—the cells used in the HSCT did not lack CCR5, and the procedure was performed with a reduced intensity chemotherapy regimen. This approach—if successful—may sidestep some of the issues of HSCT toxicity related to chemotherapy, and the relatively low frequency of *ccr5*Δ*32* homozygous donors that has plagued attempts to repeat the success of the first cure. However, as the Boston patients have only been off ART for a short period of time, and prior HSCTs performed in

HIV-infected individuals who were subsequently placed on ART have not resulted in a cure (118,119), it remains to be seen whether the transplants have actually cured their infection or simply delayed the return of the virus. Of note, the patient who received a transplant using *ccr5*Δ32 cells and both Boston patients who survived allogeneic HSCT with *ccr5* wild-type cells were all *ccr5*Δ32 heterozygotes at baseline. At this time, however, it is not known whether their baseline *ccr5*Δ32 heterozygosity could have provided any added benefit in the form of a smaller viral reservoir, or even a less severe course of their GVHD. [92,93]

Although ART does not fully restore health, its introduction has resulted in a near-normalization of the life expectancy of HIV-infected individuals, and as such, future studies must determine whether ZFN treated cells can sufficiently control viral replication, restore normal immune function, and decrease the viral reservoir, as these features of HIV infection must be impacted if we hope to achieve a cure(48,120)—functional or otherwise—in HIV infected individuals.

In the second part of my thesis, I set out to ask the following three questions: (1) Do double strand breaks generated by a single designer nuclease at on and off-target loci result in the formation of translocations between these breaks? (2) Can these translocations be detected using Translocation Capture sequencing (TCseq)? (3) Will the characterization of the off-target sites involved in these translocations further our understanding of the factors governing the specificity of these nucleases?

While this study is still ongoing, we were able to address our first question. Using nested PCR, we were able to identify rearrangements involving double strand breaks generated by a CRISPR/Cas nuclease at its on- and off-target loci. We hope that our TCseq analysis will provide much deeper insight into the genome-wide distribution of designer nuclease induced rearrangements. However, the identification of these rearrangements highlights two extremely important issues facing the field of precision genome engineering—the paucity of functional readouts of off-target activity, and the need to define an 'acceptable' level of off-target activity, if indeed such a thing exists.

While several studies have identified sites of designer nuclease off-target activity, few have gone on to address the consequences of these off-target effects. This is in part due to the fact that a large proportion of these off-target modifications occur either in non-coding regions, in genes of unknown function, or occur at a relatively low frequency compared to the off-target site. And while a low level of modification in a non-coding sequence may be of little consequence, our work has shown that a single designer nuclease can result in translocations between protein-coding genes. The identification of these translocations provides us the unique opportunity to begin to probe the biological products of off-target gene modification. Specifically, we can begin to assess the oncogenic potential of designer nucleases by analyzing the rearrangements they produce and any fusion proteins that may result. Additionally, TCseq provides a rapid means of assessing the genome-wide specificity of any designer nuclease, and as such,

can be used to quickly determine whether improvements in nuclease architecture result in a decrease in genome-wide off-target activity.

The second equally important issue highlighted by this study is the need to define an acceptable safety profile for a given designer nuclease. Early studies attempted to do this by using the *in vitro* toxicity of a nuclease as a readout for its safety. However, this approach likely overlooks a myriad of potentially genotoxic mutations whose phenotypes may not result in death, but may prove quite harmful *in vivo*. The benefits of our approach are two-fold—it may provide the most comprehensive analysis of genome-wide off-target profiles for any designer nucleases tested—and these profiles are linked to translocations that can be directly tested for biological activity. This functional readout may allow for the detection of clinically actionable rearrangements that occur following the delivery of a single designer nuclease.

The field of genome engineering has huge potential for changing the way we grow crops, breed animals, and treat humans, and I hope that this thesis has yielded useful contributions to this burgeoning and exciting field.

Figures and Tables

Table 1. Gene therapy clinical trials targeting the major steps in HIV entry

Step in Entry Targeted	Transgene/Payload	Target cell	% gene marked cells at study conclusion
CD4 binding	CD4-zeta chimeric T cell receptor	Autologous CD4+ and CD8+ T cells	0.1% of PBMCs at 1 year
CD4 binding	CD4-zeta chimeric T cell receptor	Syngeneic CD4+ and CD8+ T cells	0.1-1% of CD4+ and CD8+ T cells at 1 year
CD4 binding	CD4-zeta chimeric T cell receptor	Autologous CD4+ and CD8+ T cells	0.1-10% of PBMCs at 24 weeks
CCR5 binding	shRNA targeting tat/rev, TAR decoy, CCR5 ribozyme	Autologous CD34+ Hematopoietic stem cells	0.01% of whole blood at 18 months
CCR5 binding	CCR5 specific Zinc Finger Nuclease	Autologous CD4+ T cells	Completed
CCR5 binding	CCR5 specific Zinc Finger Nuclease (dose escalation)	Autologous CD4+ T cells	Ongoing
CCR5 binding	CCR5 specific Zinc Finger Nuclease (dose escalation, with cyclophosphamide)	Autologous CD4+ T cells	Ongoing
CCR5 binding	CCR5 specific Zinc Finger Nuclease	Autologous hematopoietic stem cells	Ongoing
HIV fusion	HIV fusion inhibitory peptide maC46	Autologous CD4+ T cells	Less than 0.01% of leukocytes afted day 7

Table 2. Indel frequency at predicted off-target sites following simultaneous treatment of primary CD4+ T cells with the R5- and X4-ZFNs

						Mock treated samples			R5/X4-ZFN treated samples (MOI_500/500)					
Arrangement	Start Site	Site	Chromosome	Mismatches	Gene	Total Reads	%indels (M)		Total Reads	%indels (Z)		(Z-M)Indel%	(Z-M)/M	Conclusion
R5ZFN-L_N5_R5ZFN-R	46414544	GTCATCCTCATCCTGATAAACTGCAAAAG	chr 3	0	CCR5	6033	0.10		5987	28.04		27.94	280.98	on target
X4ZFN-L_N6_X4ZFN-R	136872909	CCACCCACAAGTCATTGGGGTAGAAGCGGTCA	chr3	0	CXCR4	7836	0.14		6937	20.87		20.73	147.70	on target
X4ZFN-L_N5_R5ZFN-R	98297881	AGtCCGCTTCTAaAGAGTAAACTGCAtAAGT	chr1	3	DPYD	12597	0.16		18538	0.18		0.02	0.12	no activity
R5ZFN-R_N5_X4ZFN-L	109228230	AaTTTTACGGTTaACAATGgAGAAGCGGTAC	chr6	3	ARMC2	13042	0.09		14217	0.07		-0.02	-0.24	no activity
X4ZFN-L_N6_R5ZFN-R	27016543	AGACCGCTTCTAtATCCATAAAaTGGAgAAGA	chr3	3		19337	0.01		20643	0.02		0.01	1.34	no activity
X4ZFN-L_N5_R5ZFN-R	21632080	GTGCCGgTTCTACCAGCTgAACaGGAAAAGC	chr14	3		14143	0.11		14975	0.15		0.04	0.39	no activity
X4ZFN-L_N6_R5ZFN-R	1850892	ATGCCaCTTCTACCGAGACcAACTGGAAATGA	chrY	3		12096	0.05		16228	0.01		-0.04	-0.75	no activity
X4ZFN-R_N6_R5ZFN-L	75770618	ACACCtCCAGGaCTACATGGcTGAGGATGAgG	chr6	4		9270	0.01		8412	0.05		0.04	3.41	no activity v
X4ZFN-R_N5_R5ZFN-L	12941646	TCACCCCCAGtTCACCCTGAgGAGtAgGACG	chr17	4		10491	0.35		12447	0.27		-0.08	-0.23	no activity
X4ZFN-R_N5_R5ZFN-L	71635397	GCACCtCCAAGTCCTCACGAgGAGGcTGAgA	chr10	4	COL13A1	12581	0.14		18794	0.87		0.72	5.06	no activity n
X4ZFN-R_N5_R5ZFN-L	110885753	ACACCCtCgGGTCCTGCAGATGAGCcTGAgT	chr13	4	COL4A1	2948	0.68		2541	0.12		-0.56	-0.83	no activity
X4ZFN-R_N6_R5ZFN-L	14425533	GCACCCCCtGGgtCTCCGGGATGAGGAgGACC	chr19	4		8665	0.23		6073	0.33		0.10	0.43	no activity
X4ZFN-R_N5_R5ZFN-L	7279610	CCACCCACtAtTCCAGCTacTGAGGATGACT	chr10	4	SFMBT2	6886	0.28		10020	0.15		-0.13	-0.46	no activity
R5ZFN-L_N6_X4ZFN-R	3238977	CcTCATCCTCAgCAGGCCTGgCTgGGGGGTGG	chr11	4		4114	4.64		3224	4.47		-0.18	-0.04	no activity n
R5ZFN-L_N6_X4ZFN-R	43748539	CaTCATCCTCATCTTCAGCcACCTGTGGGcGG	chr1	3	C1orf210	7212	0.60		4421	0.54		-0.05	-0.09	no activity
X4ZFN-R_N5_R5ZFN-L	49133212	TCACCCgCAAGTCCCAGAGtTGAGGcTGACA	chrX	3	PPP1R3F	4315	0.28		5128	0.33		0.05	0.19	no activity
R5ZFN-L_N6_X4ZFN-R	43421403	TcTCATCCTCAgCACCGTTGACCTtTGGGaGC	chr12	4		14699	0.03		15132	0.02		-0.01	-0.42	no activity
X4ZFN-R_N5_R5ZFN-L	8681224	GaACCCACAGGTCACCGAGAaGAGCATcACC	chr8	3	MFHAS1	3888	0.08		4260	0.02		-0.05	-0.70	no activity
R5ZFN-L_N6_X4ZFN-R	128765042	TcTCATtCTCATCCCCTGTGACTTGGcGGTcC	chr11	4	KCNJ5	10649	0.08		20127	0.04		-0.03	-0.40	no activity
R5ZFN-L_N6_X4ZFN-R	71871779	GGTCATCCTCAgCAGAGTGGACTaaGGGGTcC	chr17	4		14648	1.65		11586	0.47		-1.18	-0.72	no activity n
R5ZFN-R_N6_X4ZFN-R	92737093	AaTTTTGCAaTTaAAATGGGACTTGGGGGTGA	chr9	3		4238	0.17		3712	0.27		0.10	0.63	no activity
R5ZFN-R_N6_X4ZFN-R	36865579	TCTTTTAtGGTTTCGGTTGGAaTTGTGGGTGG	chr13	2	C13orf38	10186	0.01		17088	0.05		0.04	4.36	no activity v
R5ZFN-R_N5_X4ZFN-R	2796248	TCTTTTCtAGTTTCCTCTGACCTGTGGGTtG	chr6	2		9217	0.03		7702	0.04		0.01	0.20	no activity
R5ZFN-R_N6_X4ZFN-R	99258148	CCTTTTGCAGTTTCTGCTGGACaTcaGGGTGA	chr7	3	CYP3A5	7999	0.24		8135	0.25		0.01	0.04	no activity
R5ZFN-R_N5_X4ZFN-R	102430357	TCTTTTTCGGTTTTGTGTGAaTcGTGGGTGT	chr6	2	GRIK2	12758	0.29		12006	0.42		0.13	0.46	no activity n
X4ZFN-R_N6_R5ZFN-R	25532770	CCACCCCCAAGTCCCTTTTAcACTaCAAAAaT	chr3	3	RARB	11035	0.42		9146	0.31		-0.11	-0.27	no activity
R5ZFN-R_N6_X4ZFN-R	175289032	TCaTTTCaGGTcTTCTATGGACTTGTGGGTGG	chr3	3	NAALADL2	13488	0.08		16734	0.14		0.06	0.76	no activity
X4ZFN-L_N6_R5ZFN-L	103817758	ATGCCGCTTCTgCCTGCAAGATGAGaAaGACA	chr8	3		5667	0.02		6578	0.05		0.03	1.58	no activity
X4ZFN-L_N6_R5ZFN-L	141380550	AGACaGCTTCTACCATGTGGATGAGGAcaACA	chr7	3	KIAA1147	12880	0.17		9524	0.06		-0.11	-0.63	no activity
R5ZFN-L_N6_X4ZFN-L	85231512	TcTCtTCCTCATCCACTTGcTAGAAGCaGTCC	chr1	4		9306	0.03		8144	0.06		0.03	0.90	no activity
R5ZFN-L_N6_X4ZFN-L	68402483	AGTCATCCgCATCCACCAAGTAGAAGgGGaAA	chr12	3		6868	0.16		8244	0.29		0.13	0.82	no activity
R5ZFN-L_N6_X4ZFN-L	105663998	CcTCcTCCTCcTCCTCCTAGTAGAAGCtGCAC	chr7	4	FLJ23834	3485	16.24		4769	18.87		2.63	0.16	no activity li
R5ZFN-L_N6_X4ZFN-L	129253775	GGTCAgCCTCATCCCATGCGTtGAAGaGGCCT	chr10	3		13892	0.13		14547	0.34		0.21	1.60	no activity
X4ZFN-L_N6_R5ZFN-L	763557	ATcCCGCaTCTACGCCAGTGAgGAGGATGACT	chr19	3	C19orf21	9110	3.30		5441	4.30		1.00	0.30	no activity n
X4ZFN-L_N5_R5ZFN-L	42517731	AGGCCGCTTCTACTTACGGcTGAaGAgGAaG	chr8	4		3877	0.15		6052	0.08		-0.07	-0.47	no activity
R5ZFN-L_N5_X4ZFN-L	180652373	TGTCcTCCTCtTCAACTCGTtGAAaCGGTCT	chr2	4	ZNF385B	5346	0.06		8293	0.05		-0.01	-0.14	no activity
X4ZFN-L_N5_R5ZFN-L	90025463	GCACCGCTTTtTACAAGGAGAaGAGCAaGAgC	chr16	4	DEF8	9942	0.03		10837	0.06		0.03	1.14	no activity
R5ZFN-L_N5_X4ZFN-L	68589345	GGTCATGCaCcTCACTCAGTAGAAGgGGCCA	chr16	3	ZFP90	13813	0.04		12597	0.02		-0.01	-0.34	no activity
X4ZFN-L_N6_R5ZFN-L	30834644	AGACCaCTTgaACTAATAGGAgGAGGATGACG	chr2	4	LCLAT1	8239	1.27		8538	0.62		-0.65	-0.51	no activity
X4ZFN-L_N6_R5ZFN-L	9859397	ACGCtGCTaCTgCCTGGGGGcTGAGGATGACA	chr3	4	TTLL3	7200	0.18		5330	0.26		0.08	0.45	no activity
X4ZFN-L_N5_R5ZFN-L	125124276	TGGCaGCTTCaACCTGCAGATGAGCAaGAaC	chr3	4		13684	0.07		20006	0.02		-0.05	-0.70	no activity
X4ZFN-L_N5_R5ZFN-L	242334646	GCtCCtCTTCTACTTAGAGcaGAGGATGACT	chr2	4	FARP2	6502	0.29		9157	0.32		0.02	0.08	no activity

N5/N6 refers to the length of the spacer between the 2 members of the ZFN pair ie either 5 or 6bp
(Z-M)/M needs to be >=2 to be considereded as specific
Lowercase letters in sequence indicate base mismatches

Table 3. Sequences targeted by ZFNs TALENs and CRISPRs used in this study

Plasmid	Target gene	Target sequence
PZFN1	CCR5	ACCTGCAGCTCTCAT
PZFN2	CCR5	ATACAGTCAGTA
TAL-left	CCR5	TTCATTACACCTGCAGCT
TAL-right	CCR5	AGTATCAATTCTGGAAGA
CCR5.gRNA	CCR5	ATACAGTCAGTATCAATTC**TGG**
CXCR4.gRNA	CXCR4	TTCTACCCCAATGACTTGT**GGG**
VEGFA.T2.gRNA	VEGFA	GACCCCCTCCACCCCGCCTC**CGG**
VEGFA.T3.gRNA	VEGFA	GGTGAGTGAGTGTGTGCGTG**TGG**

Protospacer–adjacent motif (PAM) sequences are noted in bold)

Table 4. T7E1 and transcription template preparation primers

Gene target	Forward primer name	Forward primer sequence	Reverse primer name	Reverse primer sequence	Use
CCR5	R5.T7E1.F*	TTAAAAGCCAGGACGGTCAC	R5.T7E1.R*	TGTAGGGAGCCCAGAAGAGA	T7E1
CXCR4	X4.T7E1.F*	CAACCTCTACAGCAGTGTCCTCATC	X4.T7E1.R*	GGAGTGTGACAGCTTGGAGATG	T7E1
VEGFA.T2	VT2.T7E1.F*	AGAGAAGTCGAGGAAGAGAGAG	VT2.T7E1.R*	CAGCAGAAAGTTCATGGTTTCG	T7E1
VEGFA.T3	VT3.T7E1.F*	TCCAGATGGCACATTGTCAG	VT3.T7E1.R*	AGGGAGCAGGAAAGTGAGGT	T7E1
CXCR4 gRNA	X4.T7.gRNA.F**	GAAATTAATACGACTCACTATAG**TTCTACCCCAATGACTTGTG**	X4.T7.gRNA.R**	GCACCGACTCGGTGCCACTTT	T7 promoter addition

*CXCR4, VEGFA.T2 and T3 primers were previously published; CCR5 primers were obtained from Sigma Aldrich Compozr ZFN kit

**Underlined sequence is T7 promoter, Bold sequence is CXCR4 gRNA sequence

Table 5. Nuclease doses and delivery methods used in study

Cell Type	Gene target	Nuclease type	Nuclease delivery method	Nuclease dose	Nuclease amount (µg)
Jurkat	None (GFP)	mRNA	BTX electroporation	Single dose	10µg GFP
Jurkat	CCR5	ZFN mRNA	BTX electroporation	Low	2µ-left/2µg-right
Jurkat	CCR5	ZFN mRNA	BTX electroporation	Medium	10µg-left/10µg-right
Jurkat	CCR5	ZFN mRNA	BTX electroporation	High	20µg-left/20µg-right
Jurkat	CCR5	TALEN mRNA	BTX electroporation	Low	2µ-left/2µg-right
Jurkat	CCR5	TALEN mRNA	BTX electroporation	Medium	10µg-left/10µg-right
Jurkat	CCR5	TALEN mRNA	BTX electroporation	High	20µg-left/20µg-right
Jurkat	CCR5	Cas9-mRNA/CCR5 gRNA-DNA	BTX electroporation	Low	10µg Cas9/2µg gRNA
Jurkat	CCR5	Cas9-mRNA/CCR5 gRNA-DNA	BTX electroporation	High	10µg Cas9/4µg gRNA
Jurkat	CCR5 & CXCR4	Cas9-mRNA/CCR5 gRNA-DNA/CXCR4 gRNA-mRNA	BTX electroporation	Single dose	20µg Cas9/2µg CCR5 gRNA/2µg CXCR4 gRNA
293T	None (GFP)	DNA	Lipofectamine 2000	Single dose	10µg GFP
293T	VEGFA.T2	Cas9-DNA/gRNA-DNA	Lipofectamine 2000	Single dose	5µg Cas9/5µg gRNA
293T	VEGFA.T3	Cas9-DNA/gRNA-DNA	Lipofectamine 2000	Single dose	5µg Cas9/5µg gRNA

Table 6. Gene-specific amplification primers used in sequencing library preparation

Primer name	Primer sequence	Notes
CCR5 primer F1	GTGGCTGTGTTTGCGTCTCT	Round 1 PCR
CCR5 primer F2.1	CAAGCAGAAGACGGCATACGAGATATCACGAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.2	CAAGCAGAAGACGGCATACGAGATCGATGTAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.3	CAAGCAGAAGACGGCATACGAGATTTAGGCAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.4	CAAGCAGAAGACGGCATACGAGATTGACCAAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.5	CAAGCAGAAGACGGCATACGAGATACAGTGAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.6	CAAGCAGAAGACGGCATACGAGATGCCAATAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.7	CAAGCAGAAGACGGCATACGAGATCAGATCAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.8	CAAGCAGAAGACGGCATACGAGATACTTGAAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.9	CAAGCAGAAGACGGCATACGAGATGATCAGAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
CCR5 primer F2.10	CAAGCAGAAGACGGCATACGAGATTAGCTTAGTCAGTCAGCCTCCCAGGAATCATCT TTACCA	Round 2 PCR
VT2.T7E1.F	AGAGAAGTCGAGGAAGAGAGAG	Round 1 PCR
VEGFA.T2.F2.illumina.1	CAAGCAGAAGACGGCATACGAGATTCCCTTGTCTCCAGTCAGTCAGCCCCCCAGCTA CCACCTCCT	Round 2 PCR
VEGFA.T2.F2.illumina.2	CAAGCAGAAGACGGCATACGAGATACGAGACTGATTAGTCAGTCAGCCCCCCAGCTA CCACCTCCT	Round 2 PCR
VT3.T7E1.F	TCCAGATGGCACATTGTCAG	Round 1 PCR
VEGFA.T3.F2.illumina.1	CAAGCAGAAGACGGCATACGAGATACGAGACTGATTAGTCAGTCAGCCGCAGACGG CAGTCACTAGG	Round 2 PCR
VEGFA.T3.F2.illumina.2	CAAGCAGAAGACGGCATACGAGATGCTGTACGGATTAGTCAGTCAGCCGCAGACGG CAGTCACTAGG	Round 2 PCR

Table 7. Linker amplification primers used in sequencing library preparation

Primer name	Primer sequence	Notes
L1 PCR1	CGCACCCAGTATTCAACAGGTA	Round 1 PCR
L2 PCR1	GTTTGAGATGCCTACGCCATCC	Round 1 PCR
L3 PCR1	TGGTCGGCGAACAATAGTGGTT	Round 1 PCR
L4 PCR1	TTCAGGAGGTCACTTCGCACAT	Round 1 PCR
L6 PCR1	TAGACCGCTCAGAGGTCATACT	Round 1 PCR
L7 PCR1	CATCGTCGACACACGTGATGAC	Round 1 PCR
L8 PCR1	TATGCGGGACAGGTAATACGCG	Round 1 PCR
L9 PCR1	GGAATCTATGTAGCAGGTCGCT	Round 1 PCR
L10 PCR1	CGCTTTGAGCTATGAACCCTAT	Round 1 PCR
L11 PCR1	AATGCGACACGCATCCTGATTT	Round 1 PCR
L12 PCR1	ATTGAAGGATCCGCCTCTCTAT	Round 1 PCR
L13 PCR1	GAGTCGAATGGTGTATACCTCA	Round 1 PCR
L14 PCR1	TTATTGCGGTAGTGAGGAAGGT	Round 1 PCR
L15 PCR1	GAATCCAGTAAGATCCGTGTGT	Round 1 PCR
L1 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTGTAGTTCCTCGGATCATGTCA	Round 2 PCR
L2 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTTCCGGTCTCTCATTAACTGGA	Round 2 PCR
L3 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTGTTGAACGGACAGATTAGTGC	Round 2 PCR
L4 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTCATTGCTTCTTCCCACTAGAG	Round 2 PCR
L6 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTACTTGCACTTCTGACCTAGCT	Round 2 PCR
L7 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTGACGAGTCAGTCCTACTAAAG	Round 2 PCR
L8 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTACGCGAGCCAGACTCCATATT	Round 2 PCR

L9 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTTCGCT AGAGTACGGCCTTGAA	Round 2 PCR
L10 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTTATTG AGAGAGGGAAAGAGGC	Round 2 PCR
L11 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTTTTCG GGCCTGATTTACTTCG	Round 2 PCR
L12 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTTATTG TTTGAAGGGACGCACG	Round 2 PCR
L13 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTCTCAC CCGTTCTGGAGACTT	Round 2 PCR
L14 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTAGGTC GAGTCTTGGGTAGGT	Round 2 PCR
L15 PCR2 primer	AATGATACGGCGACCACCGAGATCTACACCAGGACTGACGCTATGGTAATTGTTGTGT CTGCTTCCGCATCAGT	Round 2 PCR

Table 8. PCR Primers used for nested PCR to detect rearrangements.

Primer name	Primer sequence
VEGFA.T2.F PCR1	AGAGAAGTCGAGGAAGAGAGAG
VEGFA.T3.F PCR1	TCCAGATGGCACATTGTCAG
VEGFA.T2.F PCR2	ACAGGGGCAAAGTGAGTGAC
VEGFA.T3.F PCR2	GAAGCAACTCCAGTCCCAAA
VEGFA.OT2-1.R PCR1	TGCTCTGGATAAAGCACAAA
VEGFA.OT2-1.R PCR2	ACTGATCGATGATGGCCTATGGGT
VEGFA.OT2-2.R PCR1	TCCTGTCACAATTCCCTGAA
VEGFA.OT2-2.R PCR2	GCAGCCTATTGTCTCCTGGT
VEGFA.OT3-2.R PCR1	GGTTTCTTCCGGGATTTGTA
VEGFA.OT3-2.R PCR2	TACCCGGGCCGTCTGTTAGA
VEGFA.OT3-18.R PCR1	CACTGAAGCAGAGAGTAGAATGG
VEGFA.OT3-18.R PCR2	GTTGCCTGGGGATGGGGTAT

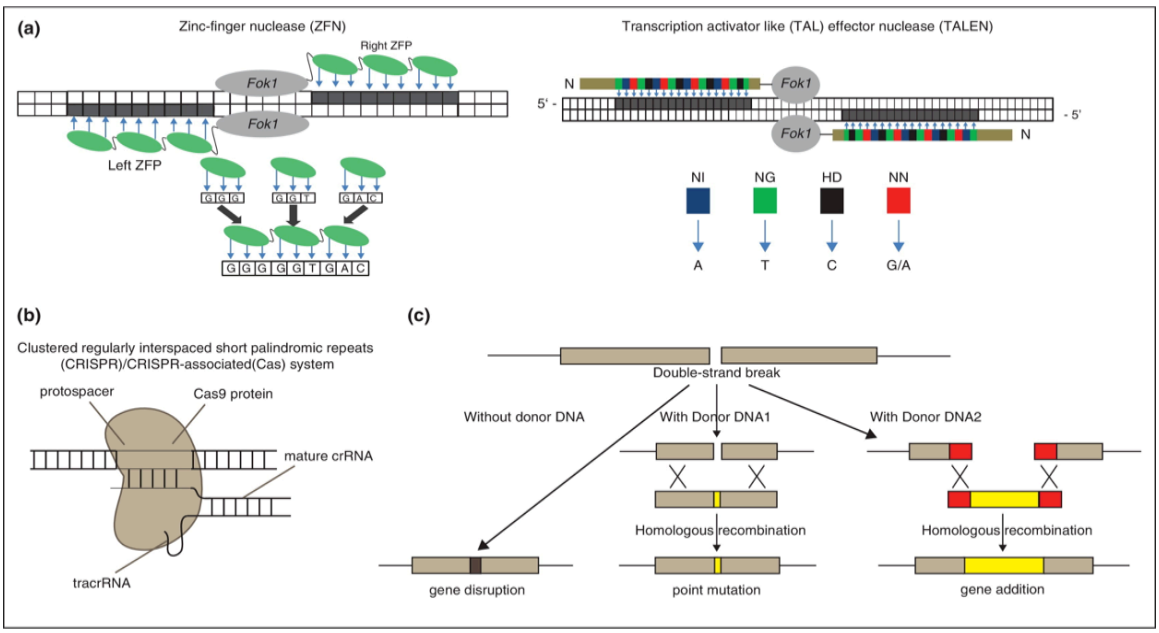


Figure 1. Designer nuclease mechanism of action. (A) Architecture of zinc finger nucleases (ZFNs) and Transcription activator-like effector nucleases (TALENs). In a ZFN, each ZFP binds 3-4 base pairs of DNA, while each repeat divariable residue (RVD; colored squares, top right panel) binds a single nucleotide. (B) Schematic depicts binding of a Cas9/crRNA/tracrRNA complex to a target sequence (protospacer) (C) Depiction of the events following the introduction of a double strand break which result in genomic alterations—gene disruption following incorrect repair by non-homologous end-joining, gene correction or gene substitution following repair by homologous recombination. Image adapted from Current Opinion in Plant Biology, Vol 16/ Issue 2, Hao Chen,Yongjun Lin, Promise and issues of genetically modified crops, 255-260 Copyright (2013), with permission from Elsevier.

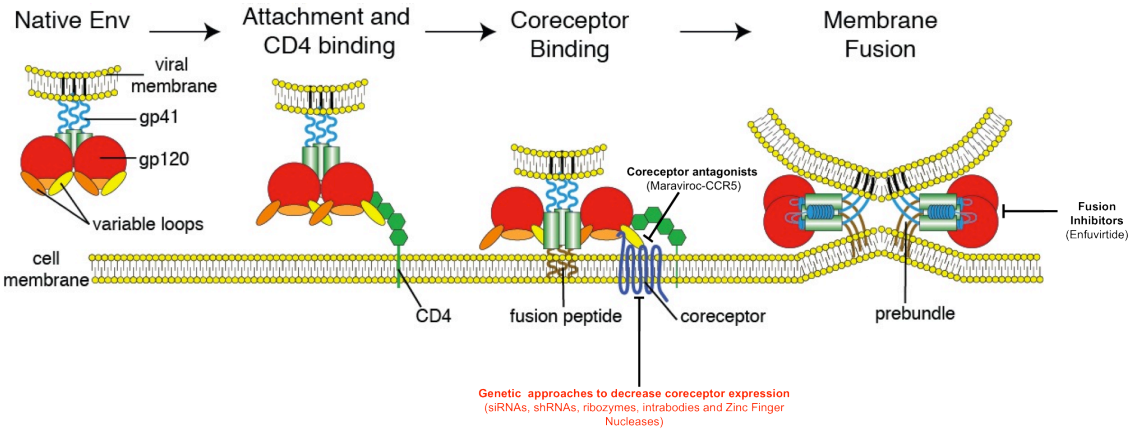


Figure 2. The HIV Entry Process involves the CD4 receptor, and CCR5 or CXCR4 coreceptors. The figure below outlines a model for HIV Entry. The entry process begins with binding of gp120 (red) to its primary cellular receptor CD4 (green). CD4 binding results in conformational changes that allow binding of gp120 to the coreceptor—either CCR5 or CXCR4. Coreceptor binding results in triggering of the fusion machinery and formation of the six-helix bundle required to drive fusion of the viral and host cell membranes. Also pictured are the two main steps that have been successfully targeted (coreceptor binding and viral fusion – approved therapeutics appear in parentheses) along with the primary target of most genetic therapies aimed at preventing HIV entry (in red). There are no approved therapies in this latter group. Adapted from Antiviral Research Vol 85, Tilton J.C. and Doms R.W, "Entry inhibitors in the treatment of HIV-1 infection," 91-100, Copyright 2009, with permission from Elsevier.

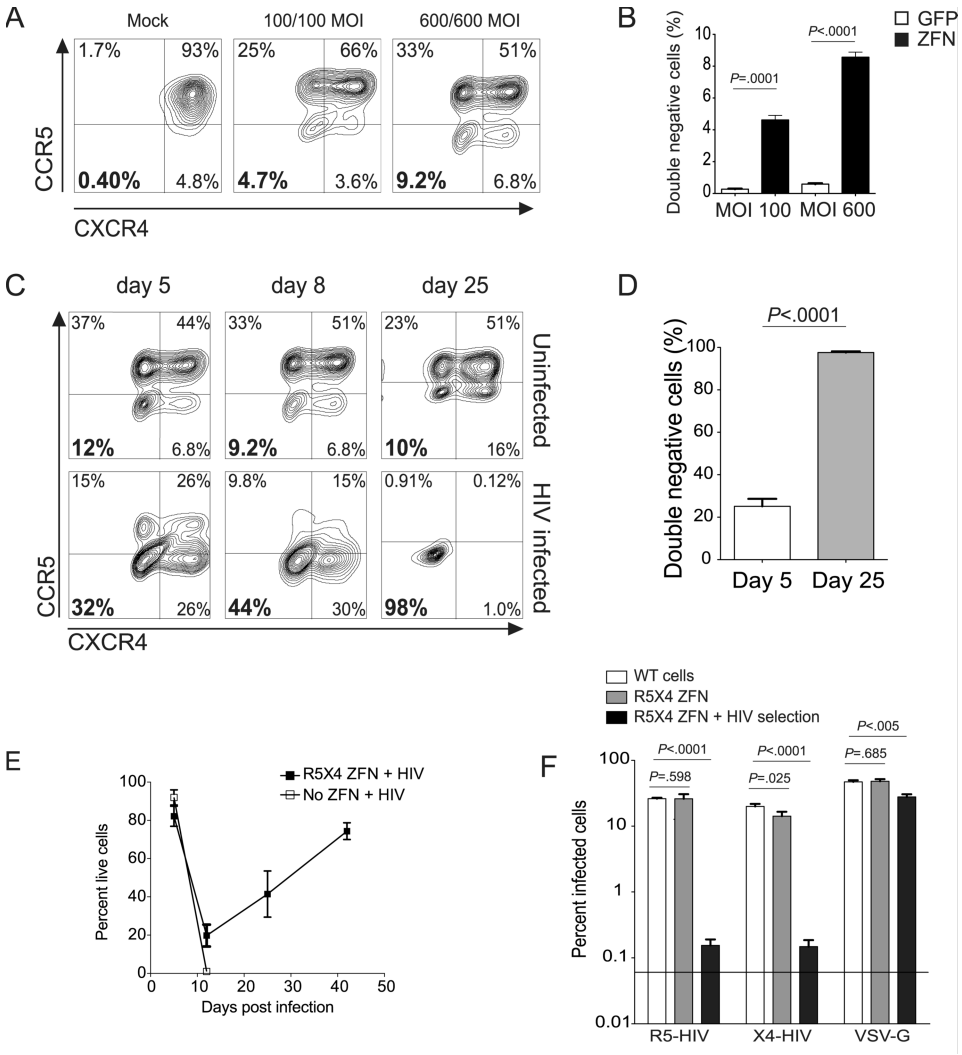


Figure 3 – Simultaneous ZFN modification of *ccr5* and *cxcr4* protects SupT1-R5 T cells from infection with viruses that use either CCR5 or CXCR4. (A) Surface expression of CCR5 (R5) and CXCR4 (X4) on SupT1-R5 cells following delivery of increasing multiplicity of infections (MOI) of the Ad R5- and X4-ZFNs. Percent of cells lacking both coreceptors is labeled in red. (B) Proportion of cells lacking both surface R5 and X4 (double negative cells) following simultaneous treatment with the R5 (MOI 600) and X4 (MOI 600) ZFNs as measured by FACS. (C) Dual ZFN treated cells were challenged with a mix of R5- and X4-using HIV and surface expression of R5 and X4 was measured over time in infected and uninfected R5/X4-ZFN treated cells. (D) The proportion of double negative cells following R5/X4-ZFN treatment and subsequent HIV infection was measured by FACS five and 25 days post-infection. (E) Cell viability after infection of mock (no ZFN) or R5/X4-ZFN treated supT1-R5 cells with a mix of R5- and X4-HIV. Viability was measured by FACS following treatment with a viability dye). (F) Dual-ZFN treated supT1-R5 cells previously challenged with HIV and no longer expressing either coreceptor (black bars) as shown in panel C, were re-challenged with either R5-HIV, X4-HIV or VSV-G-HIV pseudoviruses expressing GFP. HIV pseudovirus re-challenge of previously HIV-selected double negative cells resulted in 170-fold and 92-fold decreases in infection by R5 and X4-HIV respectively, whereas VSV-G pseudovirus infection was decreased only ~1.7-fold. All graphs represent the mean (± standard error of the mean (SEM)) of four independent ZFN treatments and four independent infection experiments; *P*-values calculated using student's *t*-test.

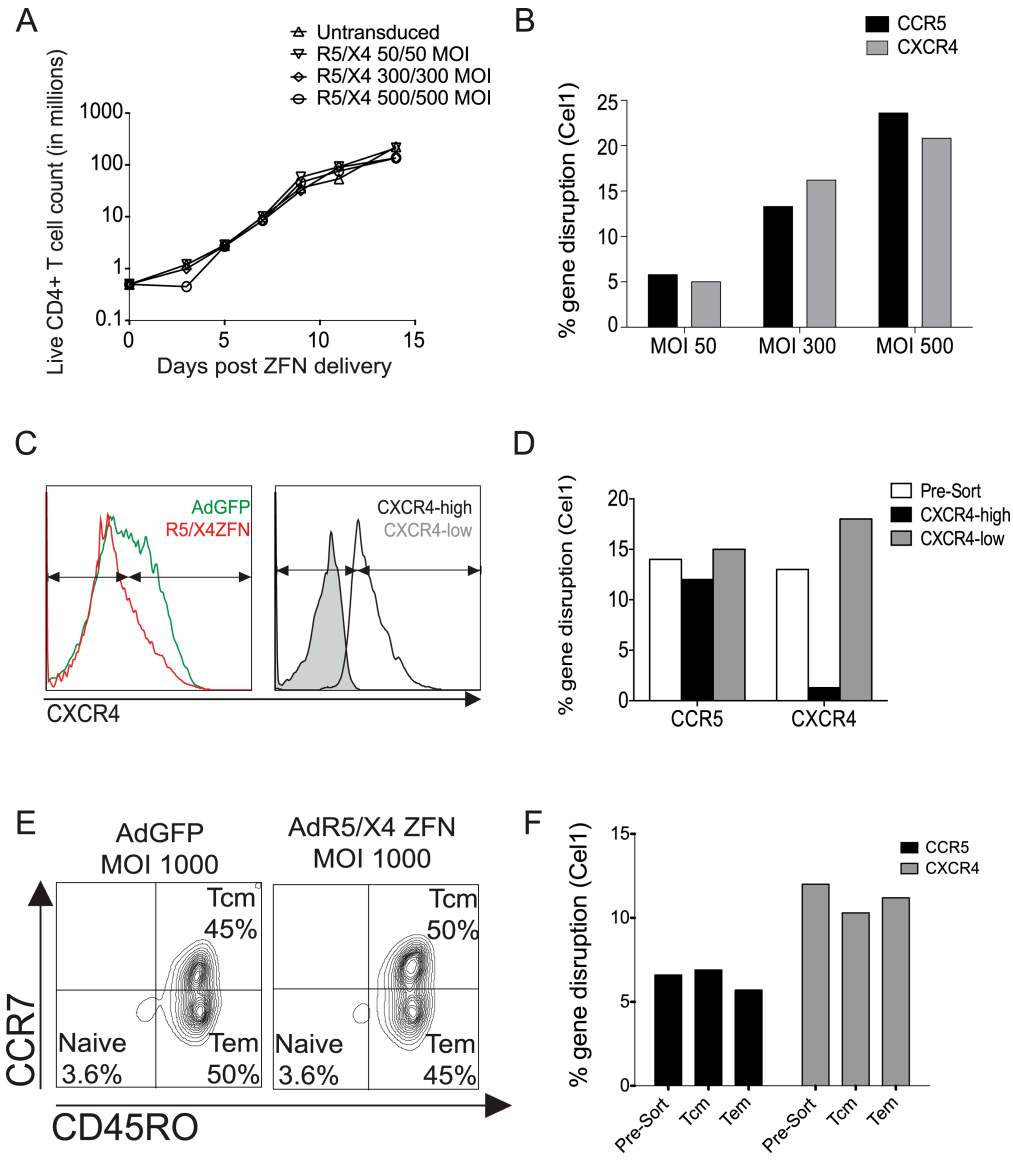


Figure 4 –ZFNs simultaneously disrupt *ccr5* and *cxcr4* in primary human CD4+ T cells. (A) Dual ZFN treatment does not result in a significant growth defect as determined by live primary CD4+ T cell count following simultaneous delivery of increasing amounts of the R5 and X4-ZFNs. Data is from one of three independent experiments. (B) *ccr5* and *cxcr4* gene modification increases in a dose-dependent manner following R5/X4-ZFN treatment, as measured by the Cel1 assay 7-10 days post-ZFN treatment. Data is from one of three independent experiments. (C) Primary CD4+ T cells were sorted by CXCR4 expression into CXCR4 high and CXCR4 low populations (left) following simultaneous R5/X4-ZFN treatment. Successful sorting was confirmed by repeat FACS (right panel). Arrows indicate gating strategy. (D) *ccr5* and cxcr4 gene disruption was measured by cel1 following R5/X4-ZFN treatment and subsequent sorting by FACS into CXCR4-high and low populations. *ccr5* disruption is similar in the X4 high and X4 low populations suggesting *ccr5* and *cxcr4* disruption are not occurring in mutually exclusive cells. Data is from one of three independent experiments. (E) CD4+ T cells treated with equal MOIs of the AdGFP or the R5- and X4-ZFNs were stained with the memory markers CCR7 and CD45RO. (Tcm; CD45RO+CCR7-) and effector memory T cells (Tem; CD45RO+CCR7-). (F) *ccr5* and *cxcr4* gene modification in Tcm and Tem subsets was similar suggesting that long-lived Tcm cells can be efficiently rendered HIV-resistant. Data shown is from one of three independent experiments.

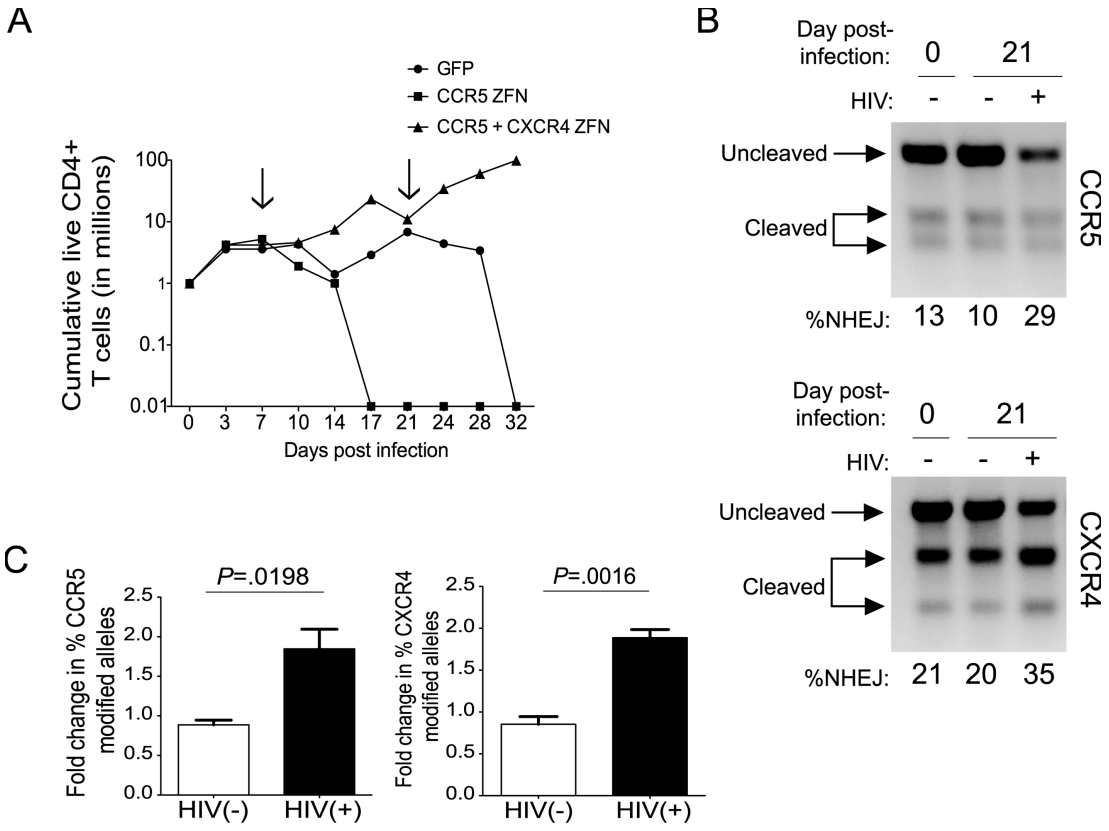


Figure 5. Primary CD4+ T cells treated with the R5 and X4-ZFNs and challenged with HIV have a survival advantage *in vitro*. (A) Primary CD4+ T cells treated with GFP, the R5-ZFN, or two vectors encoding the R5 and X4-ZFNs were challenged with a mix of two HIV strains—the R5-using virus Bal, and the X4-using virus HxB2. Dual ZFN treatment conferred a significant survival advantage in vitro compared to both the GFP and R5-ZFN control. Live cells were counted every three to four days for 32 days. Arrows indicate the time points where cells were reactivated with anti-CD3/CD28 magnetic beads. Data is from one of three independent experiments. (B) T7E1 analysis of *ccr5* (top panel) and *cxcr4* (bottom panel) gene modification reveals that HIV challenge enriches for *ccr5* and *cxcr4* disrupted cells. Analysis was performed 0 and 21 days post infection. Data is from one of three independent experiments. (C) The proportion of modified *ccr5* and *cxcr4* alleles in cells treated with both ZFNs increases approximately 2-fold in the presence or absence of HIV, as measured by the T7E1 assay. Data is shown as the mean ± SEM of three independent experiments. *P*-values calculated using student's *t*-test.

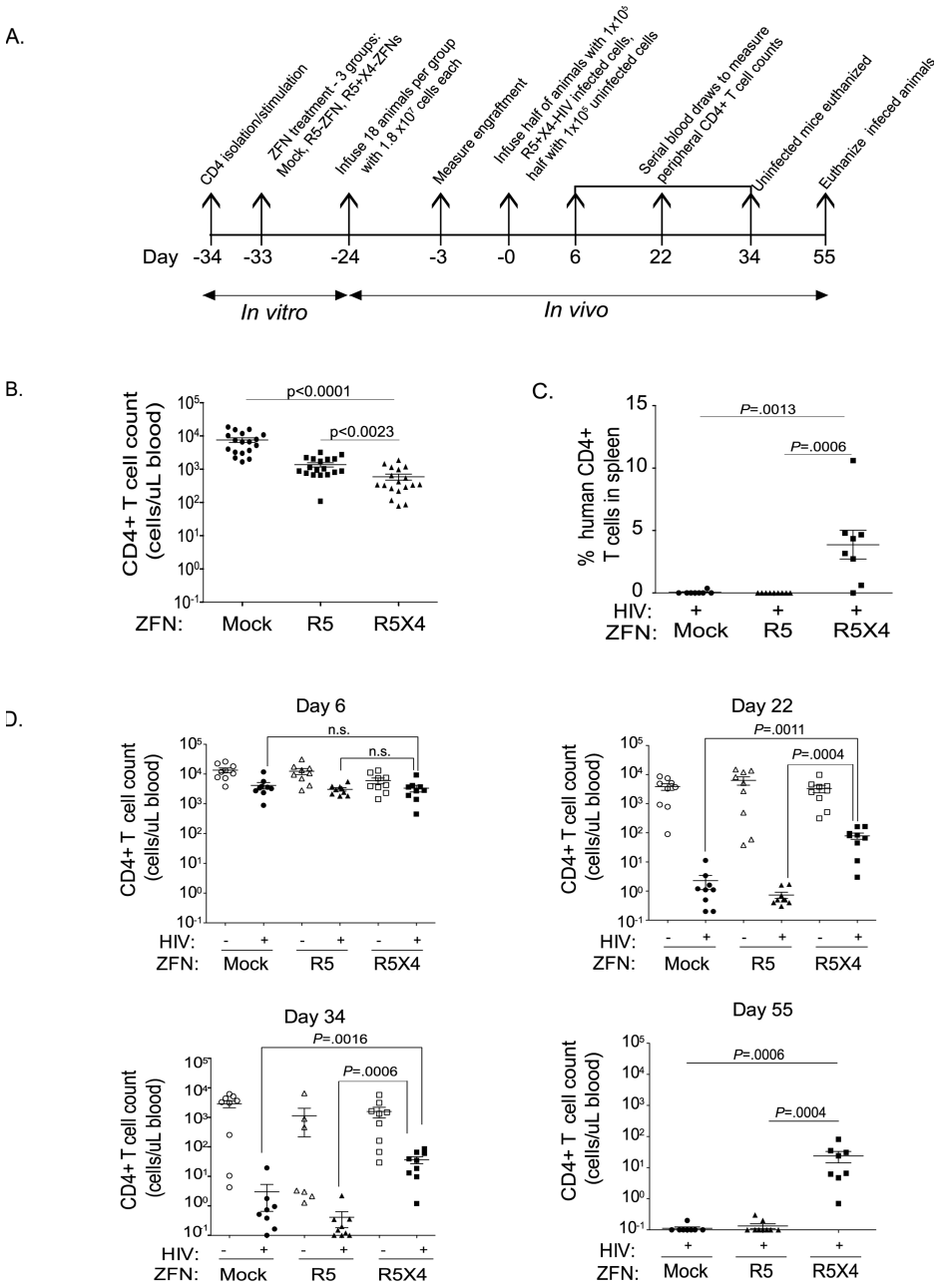


Figure 6 – Dual R5- and X4-ZFN treatment confers protection *in vivo.* (A) Experimental timeline of the *in vivo* study is shown. (B) All three treatment groups were successfully engrafted as measured 21 days post-infusion and three days prior to infection with a slight decrease in CD4+ T cell counts in ZFN groups compared to the GFP control. (C) Dual ZFN treatment preserves CD4+ T cells in the spleen compared to controls. Spleens were obtained from infected animals from all animals at the time of euthanasia. Human CD4+ T cells were defined as CD45+ CD3+ CD4+ CD8- T cells by FACS. (D) Dual ZFN treatment confers long-term protection against HIV as measured by preservation of peripheral blood CD4+ T cell counts. All uninfected animals and 1 infected animal in the R5/X4-ZFN treated group were euthanized 34 days post infection due xenogeneic graft-versus-host disease. Infected animals were followed for 55 days after infection. All statistical analyses were performed using a Mann-Whitney U test. Error bars indicate the mean ± SEM.

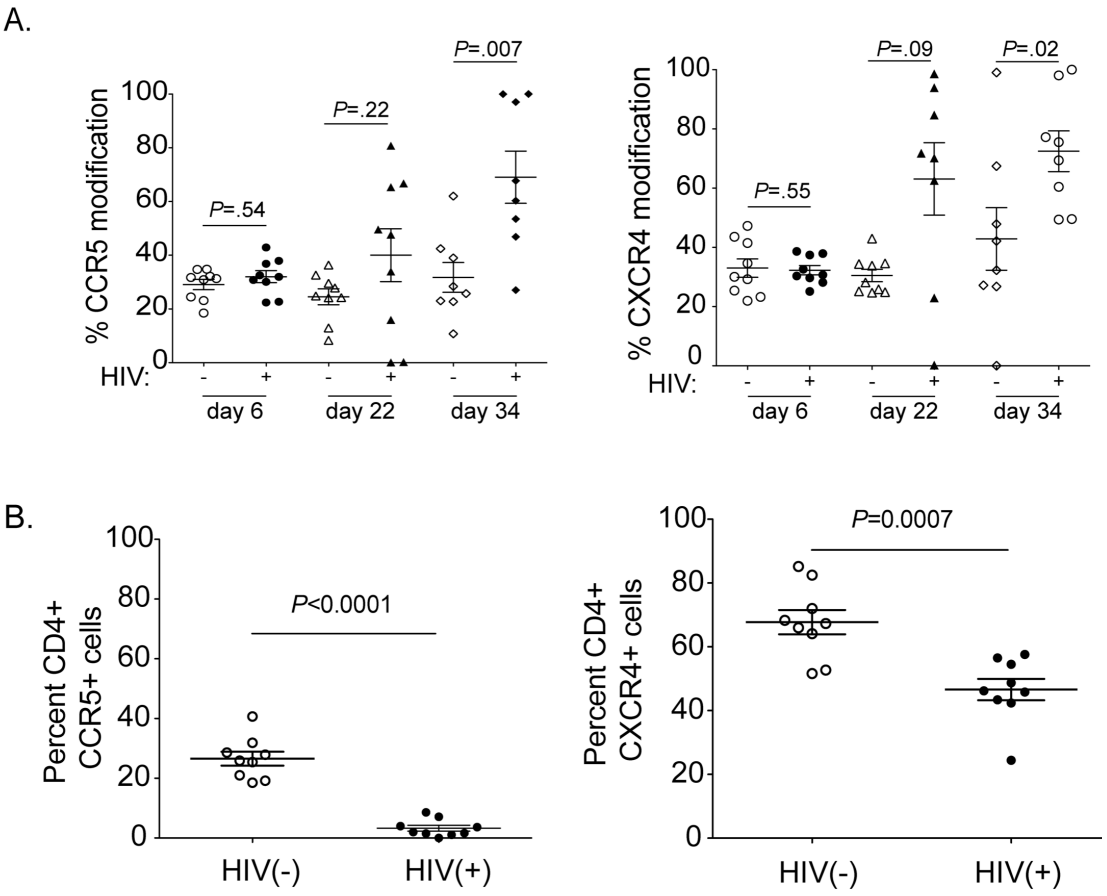
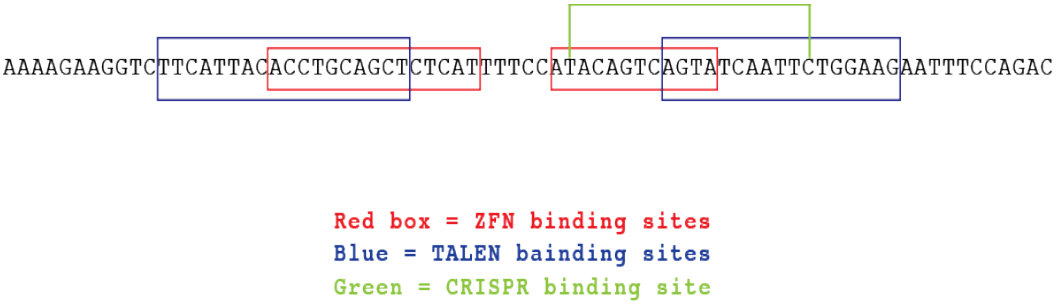
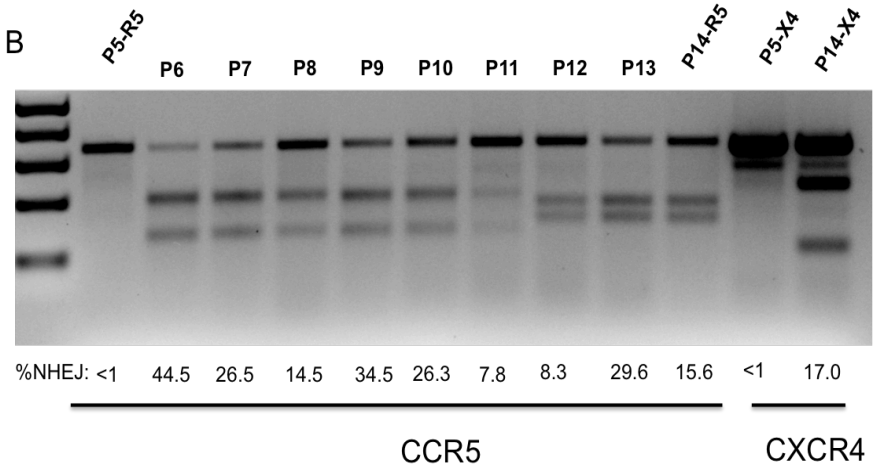


Figure 7 –Cells lacking both *ccr5* and *cxcr4* following R5/X4-ZFN treatment have a survival advantage *in vivo* in the presence of HIV.

(A) The frequency of human *ccr5* and *cxcr4*-modified genes significantly increases in the R5/X4-ZFN treatment group over the course of the *in vivo* infection. *ccr5* and *cxcr4* disruption are stable over time in the absence of HIV infection suggesting there is no significant adverse effect of dual ZFN treatment. We performed illumina deep sequencing of the R5 and X4-ZFN target sites and identified ZFN-induced mutations at these sites in 8/9 uninfected and 8/9 infected animals. We were unable to obtain sequence information from one infected animal due to limiting quantities of CD4+ T cells, and euthanized one uninfected animal following the development of GVHD, and this animal is thus excluded. (B) We stained whole blood from infected and uninfected animals in the R5/X4-ZFN treatment group 34 days post-infection with human antibodies to identify human CD4+ T cells and determine surface levels of R5 and X4. Both CCR5 and CXCR4 expression were significantly reduced on dual ZFN treated cells in the presence of HIV challenge suggesting that coreceptor negative cells have an in vivo survival advantage due to ZFN treatment.

A.

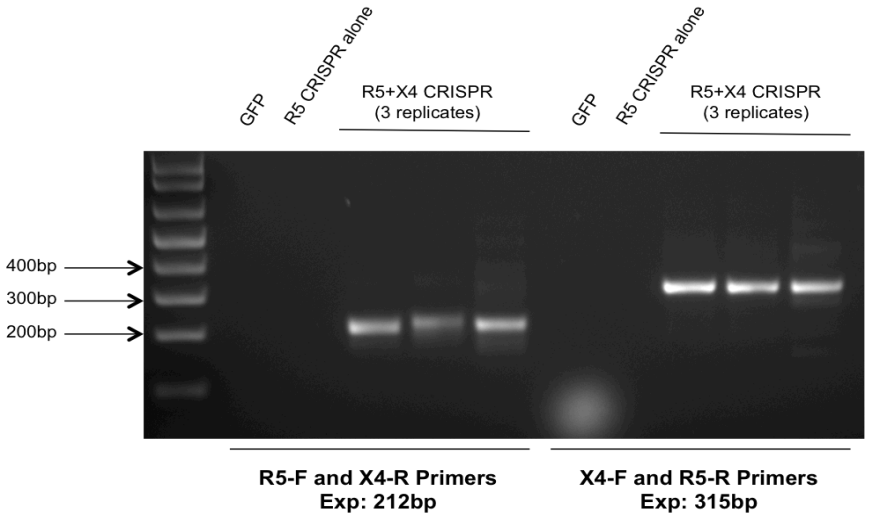




Sample ID	Description
P5	GFP
P6	ZFN High
P7	ZFN Medium
P8	ZFN low
P9	TALEN High
P10	TALEN medium
P11	TALEN low
P12	CRISPR low
P13	CRISPR high
P14-R5	CRISPR R5/X4
P14-X4	CRISPR R5/X4

Figure 8. Sequence and activity of CCR5 targeted designer nucleases (A) Sequence overlap between CCR5 ZFN binding sites (red boxes), CCR5 TALEN binding sites (Blue boxes), and CCR5 CRISPR gRNA binding site (green) (B) The CCR5 targeted nucleases area all active at their target loci. Gene modification (%NHEJ) is as noted, and samples are described in adjacent table.

A.



B.

R5X4 rearrangements

CATACAGTCAGTATCAGCTCTGTGGGTGGTTGTGTTCCAGTTTCAG

CATACAGTCAGTATCAAGTAGGGGGAAAAGTGGTTGTGTTCCAGTTTCAG

CATACAGTCAGTATCAAGTAGGGGGAAAAGTGGTTGTGTTCCAGTTTCAG

X4R5 rearrangements

TTCTACCCCAATGAC-GGGTTTCTGGAAGAATTTCCAGACATTAAA

TTCTACCCCAATGACTGACCCCCTTCTGGAAGAATTTCCAGACATTAAA

TTCTACCCCAATGACTGGGCCTAATTCTGGAAGAATTTCCAGACATTAAA

TATCTGTGACCGCCTTGCCTGC----------TTCTGGAAGAATTTCCAGACATTAAA

TTCTACCCCAATGACTGTTTCTGGCTTCTGGGGTTAGCATTCCCAGACATTAAA

TTCTACCCCAATGACTGTTTCTGGCTTCAGGGGTTAGAATTTCCAGACATTAAA

Figure 9. **Reciprocal translocations between CCR5 and CXCR4 CRISPRs following simultaneous administration** (A) Nested PCR using CCR5 and CXCR4 specific primers to identify reciprocal translocations following simultaneous delivery of CRISPRs targeting both loci. Expected band sizes are as noted (B) Confirmation of CCR5:CXCR4 translocations using sanger sequencing. Representative sequences are shown

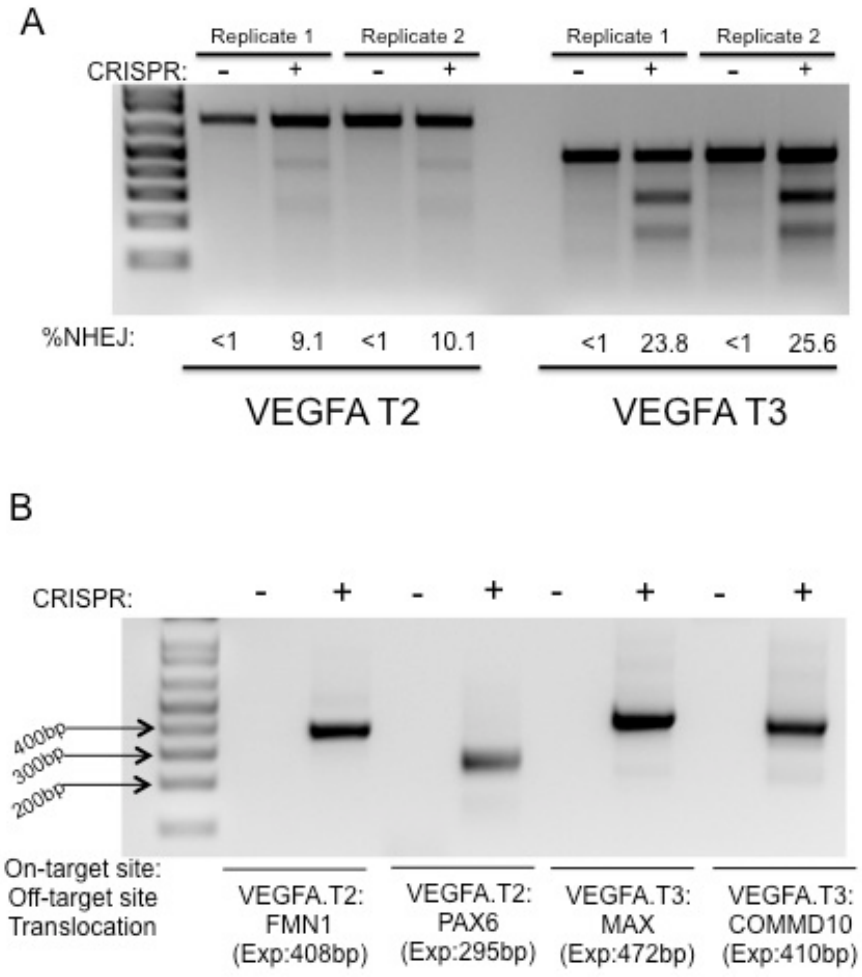


Figure 10. Identification of VEGFA CRISPR/Cas nuclease-induced rearrangements using nested PCR (A) T7 Endonuclease assay to measure gene disruption following delivery of CRISPR/Cas nucleases targeting two loci (T2 and T3) in the VEGFA gene. %NHEJ is as indicated. (B) Nested PCR on genomic DNA from 293T cells identified rearrangements between CRISPR/Cas on- and off-target loci only in cells that received the nucleases. Expected band sizes are as shown

A.

VEGFA.T2:FMN1

ccggaggcggggtggaggggtcttttggaaatgcaaatctcc

cccggcgcggggtgggggggtcttttggaaatgcaaatctgc

gtggacgcggcggcgaggggtcttttggaaatgcaaatctgc

ccggaggcggggtggaggggtcttttggaaatgcaaatctgc

cggcggcggacagtggacg-56bp-gggggtcttttggaaatgcaaatctcc

ccacctcctccccggc-75bp-gcggggtggaggggtcttttggaaatgcaaatctcc

ccacctcctccccggc-75bp-gcggggtggaggggtcttttggaaatgcaaatctcc

Gccgcgggcaggggccggagcccgc-25bp/20bp-tcatgtcacatattcctctgctc

B.

VEGFA.T2:PAX6

caggggccggagcccgcgcccgg-16bp/33bp-ccaggaatattatttataaccagga

caggggccggagcccg-23bp-ggcggggtggaggggcccctaggagcgccttggtggg

caggggccggagcccg-23bp-ggcggggtggaggggcccctaggagcgccttggtggg

ggacgcggcggcgag-46bp-gcggggtggaggggcccctaggagcgccttggtggg

ggccggcggcgg-66bp-cggggtggaggggcccctaggagcgccttggtggg

C.

VEGFA.T3:MAX

gtgtgggtgagtgagtgtgtgc--3bp-cactcactcgctctctcac

gtgtgggtgagtgagtgt------7bp-cactcactcactcactcac

gtgtgggtgagt----------13bp-cactcactcactcactcac

D.

VEGFA.T3:COMMD10

acgtgtgtgt----------27bp-cgtgaggacatttaagatcta

acgtgtgtgtctgtgtgggtgagtgag--8bp-gacatttaagatcta

tgtgggtgagtgagtgtgtgc-8bp-tttaagatcta

tgtgtgggtgagtgagtgtgtg-3bp-gtgaggacatttaagatcta

Figure 11. Sanger sequencing reveals genomic rearrangements following treatment of 293T cells with a single VEGFA-targeting CRISPR/Cas nuclease (A) Sanger sequencing results showing translocations between VEGFA.T2 and off-target site FMN1. (B) Sanger sequencing results showing translocations between VEGFA.T2 and off-target site PAX6. (C) Sanger sequencing results showing translocations between VEGFA.T3 and off-target site MAX. (D) Sanger sequencing results showing translocations between VEGFA.T3 and off-target site COMMD10. VEGFA sequences are in black and off-target site sequences are in blue. Deletions are indicated as number of base pairs and may involve either or both genes, and black nucleotides indicate mutations.

BIBLIOGRAPHY

1. Pauwels K, Podevin N, Breyer D, Carroll D, Herman P. Engineering nucleases for gene targeting: safety and regulatory considerations. N Biotechnol. 2013 Jul 12.

2. Liang F, Han M, Romanienko PJ, Jasin M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc Natl Acad Sci USA. 1998 Apr 28;95(9):5172–7. PMCID: PMC20233

3. Mani M, Kandavelou K, Dy FJ, Durai S, Chandrasegaran S. Design, engineering, and characterization of zinc finger nucleases. Biochem. Biophys. Res. Commun. 2005 Sep 23;335(2):447–57.

4. Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol. 2007 Jul;25(7):778–85.

5. Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res. 2005;33(18):5978–90. PMCID: PMC1270952

6. Liu Q, Segal DJ, Ghiara JB, Barbas CF. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. Proc Natl Acad Sci USA. 1997 May 27;94(11):5525–30. PMCID: PMC20811

7. Söllü C, Pars K, Cornu TI, Thibodeau-Beganny S, Maeder ML, Joung JK,

et al. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. Nucleic Acids Res. 2010 Dec 1;38(22):8269–76. PMCID: PMC3001086

8. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu J-K, et al. Structural basis for sequence-specific recognition of DNA by TAL effectors. Science. 2012 Feb 10;335(6069):720–3.

9. Mak AN-S, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The crystal structure of TAL effector PthXo1 bound to its DNA target. Science. 2012 Feb 10;335(6069):716–9.

10. Sorek R, Lawrence CM, Wiedenheft B. CRISPR-mediated adaptive immune systems in bacteria and archaea. Annu. Rev. Biochem. 2013;82:237–66.

11. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012 Aug 17;337(6096):816–21.

12. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-Guided Human Genome Engineering via Cas9. Science. 2013 Feb 14;339(6121):823–6.

13. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. Nature Publishing Group; 2013 Jan 29;31(3):227–9.

14. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol. 2008 Jul;26(7):808–16. PMCID: PMC3422503

15. Wilen CB, Wang J, Tilton JC, Miller JC, Kim KA, Rebar EJ, et al. PLoS Pathogens: Engineering HIV-Resistant Human CD4+ T Cells with CXCR4-Specific Zinc-Finger Nucleases. PLoS Pathog. 2011 Apr;7(4):e1002020. PMCID: PMC3077364

16. Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA, et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res. 2010 Aug;38(15):e152. PMCID: PMC2926620

17. Yuan J, Wang J, Crain K, Fearns C, Kim KA, Hua KL, et al. Zinc-finger nuclease editing of human cxcr4 promotes HIV-1 CD4(+) T cell resistance and enrichment. Mol. Ther. 2012 Apr;20(4):849–59. PMCID:

PMC3321595

18. Ellis BL, Hirsch ML, Porter SN, Samulski RJ, Porteus MH. Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. Gene Ther. 2012 Jan 19.

19. Qu X, Wang P, Ding D, Li L, Wang H, Ma L, et al. Zinc-finger-nucleases mediate specific and efficient excision of HIV-1 proviral DNA from infected and latently infected human T cells. Nucleic Acids Res. 2013 Jun 26.

20. Urnov FD, Miller JC, Lee Y-L, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature. 2005 Jun 2;435(7042):646–51.

21. Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V, et al. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol. 2010 Aug;28(8):839–47. PMCID: PMC3080757

22. Li L, Krymskaya L, Wang J, Henley J, Rao A, Cao L-F, et al. Genomic Editing of the HIV-1 Coreceptor CCR5 in Adult Hematopoietic Stem and Progenitor Cells Using Zinc Finger Nucleases. Mol. Ther. 2013 Jun;21(6):1259–69. PMCID: PMC3677314

23. Jiang J, Jing Y, Cost GJ, Chiang J-C, Kolpa HJ, Cotton AM, et al. Translating dosage compensation to trisomy 21. Nature. 2013 Jul 17;500(7462):296–300.

24. Hoffmann C. The epidemiology of HIV coreceptor tropism. Eur. J. Med. Res. 2007 Oct 15;12(9):385–90.

25. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Meth. 2011 Aug 7;8(9):765–70.

26. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppel C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol. Nature Publishing Group; 2011 Aug 7;29(9):816–23.

27. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N. Engl. J. Med. 1998 Mar 26;338(13):853–60.

28. Mocroft A, Vella S, Benfield TL, Chiesi A, Miller V, Gargalianos P, et al. Changing patterns of mortality across Europe in patients infected with HIV-1. EuroSIDA Study Group. The Lancet. 1998 Nov 28;352(9142):1725–30.

29. Forrest DM, Seminari E, Hogg RS, Yip B, Raboud J, Lawson L, et al. The incidence and spectrum of AIDS-defining illnesses in persons treated with antiretroviral drugs. Clin. Infect. Dis. 1998 Dec;27(6):1379–85.

30. Kaplan JE, Hanson D, Dworkin MS, Frederick T, Bertolli J, Lindegren ML, et al. Epidemiology of Human Immunodeficiency Virus–Associated Opportunistic Infections in the United States in the Era of Highly Active Antiretroviral Therapy. Clinical Infectious Diseases. 2000 Apr;30(s1):S5–S14.

31. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med. 2010 Dec 30;363(27):2587–99. PMCID: PMC3079639

32. Van Damme L, Corneli A, Ahmed K, Agot K, Lombaard J, Kapiga S, et al. Preexposure prophylaxis for HIV infection among African women. N Engl J Med. 2012 Aug 2;367(5):411–22. PMCID: PMC3687217

33. Thigpen MC, Kebaabetswe PM, Paxton LA, Smith DK, Rose CE, Segolodi TM, et al. Antiretroviral Preexposure Prophylaxis for Heterosexual HIV Transmission in Botswana. N Engl J Med. 2012 Aug 2;367(5):423–34.

34. Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N Engl J Med. 2012 Aug 2;367(5):399–410. PMCID: PMC3770474

35. Deeks SG. HIV infection, inflammation, immunosenescence, and aging. Annu. Rev. Med. 2011 Feb 18;62:141–55.

36. Triant VA, Lee H, Hadigan C, Grinspoon SK. Increased acute myocardial infarction rates and cardiovascular risk factors among patients with human immunodeficiency virus disease. J. Clin. Endocrinol. Metab. 2007 Jul;92(7):2506–12. PMCID: PMC2763385

37. Odden MC, Scherzer R, Bacchetti P, Szczech LA, Sidney S, Grunfeld C, et al. Cystatin C level as a marker of kidney function in human immunodeficiency virus infection: the FRAM study. Arch Intern Med. 2007 Nov 12;167(20):2213–9. PMCID: PMC3189482

38. Neuhaus J, Jacobs DR, Baker JV, Calmy A, Duprez D, La Rosa A, et al. Markers of inflammation, coagulation, and renal function are elevated in adults with HIV infection. J INFECT DIS. 2010 Jun 15;201(12):1788–95. PMCID: PMC2872049

39. Kirk GD, Merlo C, O' Driscoll P, Mehta SH, Galai N, Vlahov D, et al. HIV infection is associated with an increased risk for lung cancer, independent of smoking. Clin. Infect. Dis. 2007 Jul 1;45(1):103–10.

40. Monforte AD, Abrams D, Pradier C, Weber R, Reiss P, Bonnet F, et al. HIV-induced immunodeficiency and mortality from AIDS-defining and non-AIDS-defining malignancies. AIDS. 2008 Oct 18;22(16):2143–53. PMCID: PMC2715844

41. Chun TW, Finzi D, Margolick J, Chadwick K, Schwartz D, Siliciano RF. In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. Nat Med. 1995 Dec;1(12):1284–90.

42. Chun TW, Carruth L, Finzi D, Shen X, DiGiuseppe JA, Taylor H, et al. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature. 1997 May 8;387(6629):183–8.

43. Finzi D. Identification of a Reservoir for HIV-1 in Patients on Highly Active Antiretroviral Therapy. Science. 1997 Nov 14;278(5341):1295–300.

44. Wong JK. Recovery of Replication-Competent HIV Despite Prolonged Suppression of Plasma Viremia. Science. 1997 Nov 14;278(5341):1291–5.

45. Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, et al. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. Nat Med. 1999 May;5(5):512–7.

46. Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, et al. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. Nat Med. 2003 Jun;9(6):727–8.

47. Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JA, Baseler M, et al. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc Natl Acad Sci USA. 1997 Nov 25;94(24):13193–7. PMCID: PMC24285

48. Siliciano JD, Siliciano RF. HIV-1 eradication strategies: design and assessment. Curr Opin HIV AIDS. 2013 Jul;8(4):318–25.

49. Archin NM, Liberty AL, Kashuba AD, Choudhary SK, Kuruc JD, Crooks AM, et al. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. Nature. 2012 Jul 26;487(7408):482–5. PMCID: PMC3704185

50. Shan L, Deng K, Shroff NS, Durand CM, Rabi SA, Yang H-C, et al. Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. Immunity. 2012 Mar 23;36(3):491–501. PMCID: PMC3501645

51. Ho Y-C, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DIS, et al. Replication-Competent Noninduced Proviruses in the Latent Reservoir Increase Barrier to HIV-1 Cure. Cell. 2013 Oct;155(3):540–51.

52. Eriksson S, Graf EH, Dahl V, Strain MC, Yukl SA, Lysenko ES, et al. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. PLoS Pathog. 2013 Feb;9(2):e1003174. PMCID: PMC3573107

53. Bebenek K, Abbotts J, Wilson SH, Kunkel TA. Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. J. Biol. Chem. 1993 May 15;268(14):10324–34.

54. Bebenek K, Abbotts J, Roberts JD, Wilson SH, Kunkel TA. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J. Biol. Chem. 1989 Oct 5;264(28):16948–56.

55. Bebenek K, Roberts JD, Kunkel TA. The effects of dNTP pool imbalances on frameshift fidelity during DNA replication. J. Biol. Chem. 1992 Feb 25;267(6):3589–96.

56. Inoue M, Hoxie JA, Reddy MV, Srinivasan A, Reddy EP. Mechanisms associated with the generation of biologically active human immunodeficiency virus type 1 particles from defective proviruses. Proceedings of the National Academy of Sciences. 1991 Mar 15;88(6):2278–82.

57. Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD, Dougherty JP. High Rate of Recombination throughout the Human Immunodeficiency Virus Type 1 Genome. Journal of Virology. 2000 Feb 1;74(3):1234–40.

58. Zhuang J, Jetzt AE, Sun G, Yu H, Klarmann G, Ron Y, et al. Human Immunodeficiency Virus Type 1 Recombination: Rate, Fidelity, and Putative Hot Spots. Journal of Virology. 2002 Nov 15;76(22):11273–82.

59. Hütter G, Nowak D, Mossner M, Ganepola S, Müssig A, Allers K, et al.

Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. N Engl J Med. 2009 Feb 12;360(7):692–8.

60. Allers K, Hütter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, et al. Evidence for the cure of HIV infection by CCR5{Delta}32/{Delta}32 stem cell transplantation. Blood. 2010 Dec 8.

61. Huang Y, Paxton WA, Wolinsky SM, Neumann AU, Zhang L, He T, et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. Nat Med. 1996 Nov;2(11):1240–3.

62. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell. 1996 Aug 9;86(3):367–77.

63. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, et al. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. Nature. 1996 Aug 22;382(6593):722–5.

64. Gulick RM, Lalezari J, Goodrich J, Clumeck N, DeJesus E, Horban A, et al. Maraviroc for previously treated patients with R5 HIV-1 infection. N Engl J Med. 2008 Oct 2;359(14):1429–41. PMCID: PMC3078519

65. Fätkenheuer G, Nelson M, Lazzarin A, Konourina I, Hoepelman AIM, Lampiris H, et al. Subgroup analyses of maraviroc in previously treated R5 HIV-1 infection. N Engl J Med. 2008 Oct 2;359(14):1442–55.

66. Henrich TJ, Hu Z, Li JZ, Sciaranghella G, Busch MP, Keating SM, et al. Long-term reduction in peripheral blood HIV type 1 reservoirs following reduced-intensity conditioning allogeneic stem cell transplantation. J INFECT DIS. 2013 Jun 1;207(11):1694–702. PMCID: PMC3636784

67. Roberts MR, Qin L, Zhang D, Smith DH, Tran AC, Dull TJ, et al. Targeting of human immunodeficiency virus-infected cells by CD8+ T lymphocytes armed with universal T-cell receptors. Blood. 1994 Nov 1;84(9):2878–89.

68. Yang OO, Tran AC, Kalams SA, Johnson RP, Roberts MR, Walker BD. Lysis of HIV-1-infected cells and inhibition of viral replication by universal receptor T cells. Proc Natl Acad Sci USA. 1997 Oct 14;94(21):11478–83. PMCID: PMC23511

69. Mitsuyasu RT, Anton PA, Deeks SG, Scadden DT, Connick E, Downs MT, et al. Prolonged survival and tissue trafficking following adoptive

transfer of CD4zeta gene-modified autologous CD4(+) and CD8(+) T cells in human immunodeficiency virus-infected subjects. Blood. 2000 Aug 1;96(3):785–93.

70. Walker RE, Bechtel CM, Natarajan V, Baseler M, Hege KM, Metcalf JA, et al. Long-term in vivo survival of receptor-modified syngeneic T cells in patients with human immunodeficiency virus infection. Blood. 2000 Jul 15;96(2):467–74.

71. Deeks SG, Wagner B, Anton PA, Mitsuyasu RT, Scadden DT, Huang C, et al. A phase II randomized study of HIV-specific T-cell gene therapy in subjects with undetectable plasma viremia on combination antiretroviral therapy. Mol. Ther. 2002 Jun;5(6):788–97.

72. Kiem H-P, Jerome KR, Deeks SG, McCune JM. Hematopoietic-Stem-Cell-Based Gene Therapy for HIV Disease. Stem Cell. Elsevier Inc; 2012 Feb 3;10(2):137–47.

73. Broxmeyer HE, Orschell CM, Clapp DW, Hangoc G, Cooper S, Plett PA, et al. Rapid mobilization of murine and human hematopoietic stem and progenitor cells with AMD3100, a CXCR4 antagonist. J Exp Med. 2005 Apr 18;201(8):1307–18. PMCID: PMC2213145

74. Steinberger P, Andris-Widhopf J, Bühler B, Torbett BE, Barbas CF. Functional deletion of the CCR5 receptor by intracellular immunization produces cells that are refractory to CCR5-dependent HIV-1 infection and cell fusion. Proc Natl Acad Sci USA. 2000 Jan 18;97(2):805–10. PMCID: PMC15412

75. Swan CH, Bühler B, Steinberger P, Tschan MP, Barbas CF, Torbett BE. T-cell protection and enrichment through lentiviral CCR5 intrabody gene delivery. Gene Ther. 2006 Oct;13(20):1480–92.

76. Bai J, Gorantla S, Banda N, Cagnon L, Rossi J, Akkina R. Characterization of anti-CCR5 ribozyme-transduced CD34+ hematopoietic progenitor cells in vitro and in a SCID-hu mouse model in vivo. Mol. Ther. 2000 Mar;1(3):244–54.

77. Shimizu S, Hong P, Arumugam B, Pokomo L, Boyer J, Koizumi N, et al. A highly efficient short hairpin RNA potently down-regulates CCR5 expression in systemic lymphoid organs in the hu-BLT mouse model. Blood. 2010 Feb 25;115(8):1534–44. PMCID: PMC2830759

78. Lee MTM, Coburn GA, McClure MO, Cullen BR. Inhibition of Human Immunodeficiency Virus Type 1 Replication in Primary Macrophages by Using Tat- or CCR5-Specific Small Interfering RNAs Expressed from a

Lentivirus Vector. Journal of Virology. 2003 Oct 27;77(22):11964–72.

79. Qin X-F, An DS, Chen ISY, Baltimore D. Inhibiting HIV-1 infection in human T cells by lentiviral-mediated delivery of small interfering RNA against CCR5. Proc Natl Acad Sci USA. 2003 Jan 7;100(1):183–8. PMCID: PMC140921

80. Anderson J, Akkina R. AIDS Research and Therapy | Full text | HIV-1 resistance conferred by siRNA cosuppression of CXCR4 and CCR5 coreceptors by a bispecific lentiviral vector. AIDS Research and Therapy. 2005.

81. Anderson J, Banerjea A, Akkina R. Bispecific short hairpin siRNA constructs targeted to CD4, CXCR4, and CCR5 confer HIV-1 resistance. Oligonucleotides. 2003;13(5):303–12.

82. Cordelier P, Kulkowsky JW, Ko C, Matskevitch AA, McKee HJ, Rossi JJ, et al. Protecting from R5-tropic HIV: individual and combined effectiveness of a hammerhead ribozyme and a single-chain Fv antibody that targets CCR5. Gene Ther. 2004 Aug 5;11(22):1627–37.

83. DiGiusto DL, Krishnan A, Li L, Li H, Li S, Rao A, et al. RNA-Based Gene Therapy for HIV with Lentiviral Vector-Modified CD34+ Cells in Patients Undergoing Transplantation for AIDS-Related Lymphoma. Sci Transl Med. 2010 Jun 16;2(36):36ra43–3.

84. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Meth. 2010 Dec 5;8(1):74–9.

85. Nilsson M, Ljungberg J, Richter J, Kiefer T, Magnusson M, Lieber A, et al. Development of an adenoviral vector system with adenovirus serotype 35 tropism; efficient transient gene transfer into primary malignant hematopoietic cells. J Gene Med. 2004 Jun;6(6):631–41.

86. Guschin DY, Waite AJ, Katibah GE, Miller JC, Holmes MC, Rebar EJ. A rapid and general assay for monitoring endogenous gene modification. Methods Mol. Biol. 2010;649:247–56.

87. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol. Nature Publishing Group; 2012 Apr 8;30(5):460–5.

88. Zhang H, Zhou Y, Alcock C, Kiefer T, Monie D, Siliciano J, et al. Novel single-cell-level phenotypic assay for residual drug susceptibility and reduced replication capacity of drug-resistant human immunodeficiency

virus type 1. Journal of Virology. 2004 Feb;78(4):1718–29. PMCID: PMC369469

89. Parrish NF, Wilen CB, Banks LB, Iyer SS, Pfaff JM, Salazar-Gonzalez JF, et al. Transmitted/Founder and Chronic Subtype C HIV-1 Use CD4 and CCR5 Receptors with Equal Efficiency and Are Not Inhibited by Blocking the Integrin α4β7. PLoS Pathog. Public Library of Science; 2012 May 31;8(5):e1002686. PMCID: PMC3364951

90. Yi Y, Chen W, Frank I, Cutilli J, Singh A, Starr-Spires L, et al. An unusual syncytia-inducing human immunodeficiency virus type 1 primary isolate from the central nervous system that is restricted to CXCR4, replicates efficiently in macrophages, and induces neuronal apoptosis. J. Neurovirol. 2003 Aug;9(4):432–41.

91. Brenchley JM, Silvestri G, Douek DC. Nonprogressive and progressive primate immunodeficiency lentivirus infections. Immunity. 2010 Jun 25;32(6):737–42. PMCID: PMC2904340

92. Pfaff JM, Wilen CB, Harrison JE, Demarest JF, Lee B, Doms RW, et al. HIV-1 resistance to CCR5 antagonists associated with highly efficient use of CCR5 and altered tropism on primary CD4+ T cells. Journal of Virology. 2010 Jul;84(13):6505–14. PMCID: PMC2903254

93. Pauwels K, Podevin N, Breyer D, Carroll D, Herman P. Engineering nucleases for gene targeting: safety and regulatory considerations. N Biotechnol. 2013 Jul 12.

94. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Meth. 2010 Dec 5;8(1):74–9.

95. Van Duyne R, Pedati C, Guendel I, Carpio L, Kehn-Hall K, Saifuddin M, et al. The utilization of humanized mouse models for the study of human retroviral infections. Retrovirology. 2009;6:76. PMCID: PMC2743631

96. Broder CC, Collman RG. Chemokine receptors and HIV. Journal of Leukocyte Biology. 1997.

97. Rossi JJ, June CH, Kohn DB. Genetic therapies against HIV. Nat Biotechnol. Nature Publishing Group; 2007 Dec 7;25(12):1444–54.

98. Chung S-H, Seki K, Choi B-I, Kimura KB, Ito A, Fujikado N, et al. CXC chemokine receptor 4 expressed in T cells plays an important role in the development of collagen-induced arthritis. Arthritis Res Ther. 2010;12(5):R188. PMCID: PMC2991023

99. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol. 2013 Sep;31(9):839–43. PMCID: PMC3782611

100. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting -globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res. 2013 Oct 31;41(20):9584–92.

101. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. Nature Publishing Group; 2013 Sep 1;31(9):827–32.

102. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. Nature Publishing Group; 2013 Aug 1;31(9):833–8.

103. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol. 2013 Jun 23. PMCID: PMC3773023

104. Hakim O, Resch W, Yamane A, Klein I, Kieffer-Kwon K-R, Jankovic M, et al. DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature. Nature Publishing Group; 2012 Feb 7;484(7392):69–74.

105. Oliveira TY, Resch W, Jankovic M, Casellas R, Nussenzweig MC, Klein IA. Translocation capture sequencing: a method for high throughput mapping of chromosomal rearrangements. J. Immunol. Methods. 2012 Jan 31;375(1-2):176–81. PMCID: PMC3285106

106. Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, et al. Translocation-Capture Sequencing Reveals the Extent and Nature of Chromosomal Rearrangements in B Lymphocytes. Cell. Elsevier Inc; 2011 Sep 30;147(1):95–106. PMCID: PMC3190307

107. Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho Y-J, et al. Genome-wide Translocation Sequencing Reveals Mechanisms of Chromosome Breaks and Rearrangements in B Cells. Cell. Elsevier Inc; 2011 Sep 30;147(1):107–19.

108. Zhao Y, Moon E, Carpenito C, Paulos CM, Liu X, Brennan AL, et al. Multiple Injections of Electroporated Autologous T Cells Expressing a Chimeric Antigen Receptor Mediate Regression of Human Disseminated Tumor. Cancer Research. NIH Public Access; 2010 Nov 14;70(22):9053–

61.

109. Gillet NA, Malani N, Melamed A, Gormley N, Carter R, Bentley D, et al. The host genomic environment of the provirus determines the abundance of HTLV-1-infected T-cell clones. Blood. 2011 Mar 17;117(11):3113–22. PMCID: PMC3062313

110. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the …. 2011.

111. Allers K, Hütter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, et al. Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. Blood. 2011 Mar 10;117(10):2791–9.

112. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. Science. 1996 Sep p. 1856–62.

113. Moore JP, Kuritzkes DR. A pièce de resistance: how HIV-1 escapes small molecule CCR5 inhibitors. Curr Opin HIV AIDS. 2009 Mar;4(2):118–24. PMCID: PMC2896203

114. Burnett JC, Zaia JA, Rossi JJ. Creating genetic resistance to HIV. Curr. Opin. Immunol. 2012 Oct;24(5):625–32. PMCID: PMC3478429

115. Aiuti A, Webb IJ, Bleul C, Springer T, Gutierrez-Ramos JC. The chemokine SDF-1 is a chemoattractant for human CD34+ hematopoietic progenitor cells and provides a new mechanism to explain the mobilization of CD34+ progenitors to peripheral blood. J Exp Med. 1997 Jan 6;185(1):111–20. PMCID: PMC2196104

116. Peled A, Petit I, Kollet O, Magid M, Ponomaryov T, Byk T, et al. Dependence of human stem cell engraftment and repopulation of NOD/SCID mice on CXCR4. Science. 1999 Feb 5;283(5403):845–8.

117. Weiden PL, Flournoy N, Thomas ED, Prentice R, Fefer A, Buckner CD, et al. Antileukemic effect of graft-versus-host disease in human recipients of allogeneic-marrow grafts. N. Engl. J. Med. Mass Medical Soc; 1979;300(19):1068–73.

118. Hutter G, Zaia JA. Allogeneic haematopoietic stem cell transplantation in patients with human immunodeficiency virus: the experiences of more

than 25 years. Clin Exp Immunol. 2011 Feb 8;163(3):284–95. PMCID: PMC3048611

119. Durand CM, Ambinder RF. Hematopoietic stem cell transplantation in HIV-1-infected individuals: clinical challenges and the potential for viral eradication. Curr Opin Oncol. 2013 Mar;25(2):180–6.

120. Margolis DM, Hazuda DJ. Combined approaches for HIV cure. Curr Opin HIV AIDS. 2013 May;8(3):230–5.